Bachelor's thesis

Biotechnology and Food Technology

Biotechnology

2016

Julia Bouirmane

GENETIC VARIATION INFLUENCING BODY SIZE IN PUREBRED DOGS

TURUN AMMATTIKORKEAKOULU
TURKU UNIVERSITY OF APPLIED SCIENCES

BACHELOR'S THESIS | ABSTRACT

TURKU UNIVERSITY OF APPLIED SCIENCES

Biotechnology and Food Technology | Biotechnology

Spring 2016 | 80

Instructors: Jani Pelkonen, M.Sc.; Jonas Donner P.hD. and Heidi Anderson P.hD.

Julia Bouirmane

GENETIC VARIATION INFLUENCING BODY SIZE IN PUREBRED DOGS

The dog has the greatest body size variation of all mammals. As a result of intensive breeding, it is likely that only a small number of genes control major traits in the dog, including size. In this study, the possibility of predicting the genetic size of a dog with previously identified genetic markers is evaluated. The ability to predict genetic size may have several health benefits for the dog, such as allowing preparation of special dietary and exercise plans. In addition, knowing the predicted size of the dog may help in choosing the right individual with the desired, size associated qualities for breeding.

The exceptionally large genetic data used in this study consists of 16 372 canines, including 77 wolves, 27 coyotes and 13 dingoes which provides the largest material ever used to evaluate the effects of genetic variants influencing body size.

The effects of six previously identified single-nucleotide polymorphisms associated with canine weight were evaluated by statistical methods for both dog weight and height prediction. In addition, the effects of the genetic markers associated with the breed-defining chondrodysplasia (short-leggedness) were identified.

All the examined size associated genetic markers were found to have a strong association with reduced body size, as expected based on previous research. Alleles fixed within the breed and the frequencies of the alleles in the multimarker combinations were found to predict the size range of the breed better than presence of individual marker combinations. On the other hand, marker combinations seem to predict individual size better. Only a few breeds were fixed for a single combination. Mostly marker combinations were found within a certain size scale. A marker combination containing ancestral alleles predicts the dog will be of similar size to the wolf. More rare combinations with more derived alleles may allow a finer scale prediction of dog body size even with only six genetic markers. In addition, new short-legged breeds were discovered and the association of breed-defining chondrodysplasia with the body size prediction was examined.

In conclusion, rough prediction of dog size is possible by examining only a handful of genetic markers, but for fine scale prediction, identification of additional genetic markers is required.

KEYWORDS:

Genetic variation, genotype, phenotype, genetic markers, single-nucleotide polymorphism, fibroblast growth factor, growth hormone receptor, high mobility group AT-hook, insulin-like growth factor, insulin growth factor receptor, stanniocalcin, breed-defining chondrodysplasia

OPINNÄYTETYÖ (AMK) | TIIVISTELMÄ

TURUN AMMATTIKORKEAKOULU

Bio- ja elintarviketekniikka | Biotekniikka

Kevät 2016 | 80

Ohjaajat: Lehtori Jani Pelkonen, Tohtori Jonas Donner ja Tohtori Heidi Andersson

Julia Bouirmane

KOKOON VAIKUTTAVA GENEETTINEN MUUNTELU ROTUKOIRILLA

Koiran koon variaatio on suurempi kuin minkään muun nisäkkään. Intensiivisen rotujalostuksen myötä vain pieni määrä geenejä kontrolloi näkyviäkin koiran ominaisuuksia, kuten kokoa. Tässä opinnäytetyössä tutkitaan onko koiran koon ennustaminen mahdollista aiemmin määriteltyjen kokoon liitettyjen geneettisten markkereiden avulla. Koiran geneettisen koon ennustamisella voi olla useita hyviä vaikutuksia, kuten yksilöllisen ruokavalion ja liikunnan suunnittelu. Lisäksi koiran koon ennustaminen voi auttaa oikean yksilön valinnassa kokoon liitettyjen ominaisuuksien perusteella.

Tässä tutkimuksessa käytetty ainutlaatuisen laaja tietokanta kattaa 16 372 koiraeläintä, mukaan lukien 77 sutta, 13 dingoa ja 27 kojoottia, tarjoten kaikkien aikojen suurimman materiaalin koiran pituuteen ja painoon vaikuttavien geenien tutkimuksiin.

Ennalta määritettyjen, kuuden koiran painoon liitetyn yhden emäksen monimuotoisuuden vaikutuksia koiran painoon ja pituuteen määriteltiin tilastollisten menetelmien avulla, sekä tutkittiin tappijalkaisuuden markkereiden vaikutusta koiran korkeuteen.

Kaikilla tutkituilla geneettisillä markkereilla, todettiin olevan vahva koiran kokoa pienentävä vaikutus, kuten jo aiemmissa tutkimuksissa on huomattu. Alleeliyhdistelmien todettiin ennustavan hyvin yksilöiden kokoa, toisaalta alleelien yleisyyksien todettiin ennustavan hyvin rodun kokoluokkaa. Rodun fiksoituneiden alleelien todettiin alleeliyhdistelmiä paremmin ennustavan rodun kokoa mutta vain muutaman rodun todettiin olevan fiksoitunut spesifiselle alleeliyhdistelmälle ja useimminkin alleeliyhdistelmät olivat fiksoituneet jollekin kokoluokalle. Alkukantaisista alleeleista koostuva alleeliryhmä ennustaa koiran samakokoiseksi suden kanssa. Enemmän johdettuja alleeleja sisältävillä harvinaisilla alleeliyhdistelmillä voidaan ennustaa tarkimmin koiran kokoa, jopa vain kuudella markkerilla. Lisäksi uusia, geneettisesti tappijalkaisia rotuja löydettiin ja lyhytjalkaisuuden vaikutusta kokoon tutkittiin.

Loppupäätelmänä voidaan todeta, että tutkittujen geneettisten markkereiden avulla on mahdollista karkeasti ennustaa koiran koko, mutta tarkkaan ennustukseen tarvitaan vielä muitakin geneettisiä markkereita.

ASIASANAT:

Geneettinen variaatio, geneettiset markkerit, genotyyppi, fenotyyppi, tappijalkaisuus, yhden emäksen monimuotoisuus

CONTENT

FIGURES

TABLES

LIST OF ABBREVIATIONS

AKC	American Kennel Club
BDC	Breed-defining chondrodysplasia
CFA	Chromosome number
FCI	Federation Cynologique Internationale
FGF4	Fibroblast growth factor 4
GHR1	Growth hormone receptor 1
GHR2	Growth hormone receptor 2
HMGA2	High mobility group AT-hook
IGF1	Insulin-like growth factor
IGF1R	Insulin growth factor receptor
SMAD2	SMAD family member 2
SNP	Single-nucleotide polymorphism
STC2	Stanniocalcin 2
STH	Standard breed height
STW	Standard breed weight

1 INTRODUCTION

Through the ages, the dog has been a loyal companion for man. They help people with everyday work, they bring joy to the lives of their families, and they are even believed to decrease depression. As a family member and a true friend, every dog should be entitled to a happy, healthy and a most of all, a long life.

One secret behind a healthy life is a balanced diet and moderate exercise. While there is great size variation between breeds, some variation in size occurs also within breeds, as well as in the size associated genes. The size of the dog determines the need for nutrition and exercise and while every dog is different, also every diet should be planned according to individual requirements.

5 % of the domestic dogs in the UK are underweight (Coursier et *al.* 2010) while up to one third of household dogs are overweight in both the US and Europe (Fodge 2010). Both of these conditions cause multiple health risks for the dog and may decrease its life expectancy. (Fodge 2010, German 2006, Ortega-Pachero et *al.* 2006, Smith et *al*. 2006, Glichman et *al.* 1997.)

A research that studied 48 Labrador retrievers for their lifetime came to the conclusion that the half of the group that was fed all their life with no restriction started to show signs of hip joint osteoarthritis at the age of six, while the group that was fed 25% less during their lifetime, developed hip joint osteoarthritis only at the age of 12. (Smith et *al.* 2006.)

In addition, obesity in companion pets is associated with various diseases which include metabolic abnormalities, such as insulin resistance and glucose intolerance, hypertension, dystocia, exercise intolerance, heat stroke and decreased immune functions. (German 2006.) In fact, the life expectancy of overweight dogs may be even up to 18 months shorter than the average life length of individuals within the normal weight range. (Fodge 2010.)

Nevertheless, also underweight dogs are at great risk of various diseases, such as a higher risk of the acute episodes of gastric dilatation-volvulus (Glickman et *al.* 1997) and for testicular degenerations (Ortega-Pachero et *al.* 2006.)

Canine size has also been associated with certain behavior in dogs. Undesirable behavior, such as hyperactivity, was found to be more common within the smaller body size range. Likewise, low canine height has been associated with fear and attachment. Short legged, smaller dogs, such as the bull-type or terriers are more likely to show aggressive behavior, including owner-directed aggression than larger dogs. One reason for the undesirable behavior of smaller sized dogs may possibly be the association of the size with neurological changes which may influence behavior in certain situations. Another suggested explanation for the behavior of smaller dogs may simply be the fact that the smaller dogs' owners may have better tolerance for such behavior. (McGreevy et *al.* 2013.)

By predicting the genetic size of the canine, the planning of a customized diet and exercise according to the genetic size markers is made easier and could both significantly increase life expectancy and benefit health. Also by knowing the genetic size of the dog and the possible negative behavior associated with it, proper exercise and training could be planned at an early age in order to prevent the negative effects of hyperactivity, fear or frustration, such as attention seeking or begging (McGreevy et *al.* 2013). In addition, if the character of the dog can be predicted by size, genetic size testing may help choose the puppy with the right size associated qualities for each purpose.

In addition to the benefits of dog size prediction, such as facilitated diet, exercise and training planning for better life quality as well as knowing the size associated qualities of the dogs, one may also be interested in the predicted size of the puppy for the purpose of dog shows, dog races or agility. In many dog sports the classes and groups are set according to the size of the dog. The most preferred size in dog sports is a moderate size and the most unfortunate size is the smallest or largest in the given class. By predicting the size of the puppy, choosing the right individual for the desired dog sport could prevent

eventual transfers to another class due to size limits. Also, in most breeds, the individuals that do not meet the size standards of the given breed will be discarded from the dog show, no matter how desirable the other features or appearance the canine may have (Federation Cynologique Internationale).

Because knowledge of the genetic size variation of the dog doubtlessly has many benefits, in this study the possibility of predicting the dog body size is evaluated. First, findings concerning six size associated markers previously found by Rimbault et *al*. (2013) are repeated, followed by the findings concerning the genetic markers causing short legged individuals by Parker et *al.* (2009).

The aim of repeating the previously published studies with a large genetic databank is to validate the information that has been found and tested with only a few individuals. The genetic data of over 16 000 enables for the largest study ever made on the genetics of dog size and its prediction. In addition to examining the possibility to predict the dog size, also the launching of a reliable commercial testing for canine size is examined and a dog body size report view for the MyDogDNA customers is developed according to the findings of this theses.

2 HISTORY OF DOGS

Knowing the origin of the canine may help with the genetic research but for a long time, the history of the domestic dog has been unclear. Depending on the research and its methods, the origin of the domestic dogs has varied across the continents, suggesting that the dog domestication started in Europe, Middle East and East Siberia, Asia or even Africa. (Parker et *al.* 2012.)

2.1 Molecular analysis of the canine ancestry

Many different methods have been used to reveal the history and origin of the dog. Analysis of the Y-chromosome indicated that the domestic dog originated from Southern East Asia, but in the studies on the multi-locus haplotypes of dogs, the origin of the dog was placed in the Middle East. In fact, the Middle Eastern wolves were found to share more autosomal haplotypes with the modern dog than with the other wolf populations. In addition, a haplotype closely related to the insulin growth factor (*IGF1)* mutation found in small dogs, seems to be found only in the Middle Eastern wolves, which might indicate that the origin of the smaller sized dogs is in the Middle East. (Rimbault et *al.* 2013, Ding et *al.* 2011, Gray et *al.* 2010, vonHoldt et *al.* 2010, Sutter et *al.* 2007.)

Studies on the mitochondrial genome indicate that the most likely candidate for the origin of the domestic dog is southern China. It has been suggested that the domestic dog travelled the same route across the main land of southern Asia to Australia. (Oskarsson et *al*. 2011.) The analyses of the mitochondrial DNA in both wolves and dogs also reveal that the wolf is the direct ancestor of the dog. (Vilà et *al.* 1997.)

A recent study of the single-nucleotide polymorphisms (SNPs) from 58 rural, indigenous or wild canids around the world shows that the domestic dog originated from southern East Asia 33 000 years ago (Wang et *al.* 2016). Most likely the domestic dog migrated in two directions from its place of origin, to the region of China about 10 500 years ago and to the Middle East about 15 000 years

ago. From the Middle East, the canine population again moved to Africa, Europe and to the Chinese regions. The American canine populations originated from the Chinese and the Middle Eastern populations an unknown time ago. The migration to Africa is unknown; however, the migration to Europe is estimated to have taken place about 10 000 years ago. (Wang et *al.* 2016.)

Most likely the wolf and dog diverged about 20 000-100 000 years ago, and it is estimated that the dog population grew from 4 600 to 17 500 individuals about 33 000 years ago. Most possibly this period may be considered as the beginning of the domestic dog. In fact, the southern Chinese indigenous dogs, genetically the closest to wolves, are believed to be the ancestor of all the other dog breeds. The southern Chinese indigenous dog diverged from the other dog breeds about 15 000 years ago. (Wang et *al.* 2016.)

2.2 Modern dog breeds

The dog has accompanied humans for thousands of years, but it was only 200-300 years ago that the extreme artificial selection for homogeneous canine strains, also called breeds, started (Dennis-Bryan 2012, Parker et *al.* 2012)**.** The selection for the desired dogs has been according to the skills of the dog in various utility areas, such as hunting, guarding or herding. Recently, as the companion dogs have become more popular, also the appearance has become one of the major determinants of the desired dogs. The kennel clubs have emphasized the breeding of healthy dogs which maintain the breed specific qualities (Federation Cynologique Internationale).

At the end of the 19th century Kennel clubs and awards were formed to specify the phenotype and qualities of the most desired specimen of the breeds (American Kennel Club). Nowadays, over 420 different breeds exist in the world, of which 343 are recognized by the FCI and 198 by the AKC. (American Kennel Club, Federation Cynologique Internationale, Dennis-Bryan 2012.) Many countries have their own kennel clubs, and the variety of the breeds they recognize varies, and the breed standards may be different from the others. As a result,

breeds in different continents or countries may be under a different selective pressure, slightly increasing the genetic and phenotypic heterogeneity of the breeds in total. However, as the kennel club standards may vary from club to club, crossing individuals of the same breed from different countries may be limited. (Parker 2012.)

Federation Cynologique Internationale divides breeds into ten categories based on behavior and these groups are further divided into two to ten groups based on appearance, origin or a specific usage (Federation Cynologique Internationale). AKC divides breeds into seven categories, based on their behavior (American Kennel Club).

The purpose of selection is to emphasize certain behavior or appearance for a specific purpose. Individuals with a desired phenotype are selected over multiple generations. Only dogs that have both of their parents from the same breed can be registered in that breed (American Kennel Club). However, the adverse effects of inbreeding are widely known and to avoid genetic disorders, no close relatives, such as siblings or parents should produce offspring. In addition, a dog may produce only 5% of the offspring of the total puppy population of the breed in the past five years. (Federation Cynologique Internationale.)

The downside of the development of the breeds is the loss of genetic variation, and even the loss of some genes and alleles. Within breeds with a small number of individuals, the number of different alleles is also reduced, which results in decreased diversity. (Irion et *al.* 2003.)

3 DOG SIZE

The dog has more size variation than any other mammal on the planet (Rimbault et *al*. 201, Sutter et *al.* 2007). The Chihuahua's average weight is only 1.8 kg, while the Tibetan Mastiffs may weigh even up to 140 kg, which is almost 100 times more than the smallest Chihuahuas. Most of the extreme variation in canine body size is between the breeds while the diversity within the breeds is considered small (Hoopes et *al.* 2012, Parker et *al*. 2009). However, in some breeds there may be major size differences. For instance, the female Saint Bernard may be up to 25 cm lower in height than a male, or a large Tibetan Mastiff may be over 50 kg heavier than a smaller individual (Dennis-Bryan 2012).

A great part of the inbreed diversity is due to the controlled breeding of the dog to produce the best individuals for the desired purpose. The different sized dogs are thought to be capable of different kind of work. For rabbit and fox hunting, a small size allows the possibility to follow the prey into the holes. Larger sizes may be better in guarding and protecting while herding requires a proportionate body with the right attitude and good physical condition. Nowadays, breed standards control the phenotype of the domestic dog, including the size, by putting selective pressure on each breed (Rimbault et *al.* 2012).

In addition to genetics, also nutrition and lifestyle have a major influence on the size of the canine. As the feed determines the energy, too little food will cause smaller body size in most mammals while too much food causes obesity. In fact obesity is becoming as big an issue in pet dogs as it is in humans. In the UK, only 35 % of domestic dogs are ideal weight, up to 40% of the household dogs are overweight and over 5 % are underweight (Coursier et *al.* 2010).

Genetic variation within the dog breeds allows an opportunity to emphasize the desired qualities in breeding. Character as well as the size is inheritable, however the environment has its share in determining the size of each individual. To understand the size variation of the dog one must first know the genetics behind it.

3.1 Genetic mapping in canines

Most breeds originate from a small population and have the same ancestor. As a result of intensive breeding and a few population bottlenecks, only a small number of genes control major traits in domestic dogs (Rimbault & Ostrander 2012, Vaysse et *al.* 2011), whereas in humans various genes may control a minor feature. Due to the genetics of the dog, many traits in dogs can be genetically located easily and most of the phenotypes measured in a few dogs apply to most of the individuals in that breed (Jones et *al.* 2008).

Single-nucleotide polymorphisms (SNP) are one of the key factors in the genetic mapping and research of the canine. SNPs are individual nucleotide positions in the genome that are different between individuals and many SNPs are inherited in large haplotype blocks within the chromosome (Alberts et *al.* 2008.) SNPs that are located in the introns are usually not visible in the phenotype whereas the rare SNPs that are located in the exons, might have a significant effect on the phenotype. However, mostly, there is a possibility that some SNPs may somehow be seen in the phenotype despite their location. In fact, a subset of the SNPs and the copy number variants are responsible for most of the variation in the genome and the phenotype. While some SNPs are inherited from the ancestors, others are highly mutation prone and produce diversity in the long run. (Alberts et *al.* 2008.)

The canine has 38 autosomal chromosomes of which up to 99 % have been sequenced and over 2.5 million SNPs located. (Lindblad-Toh et *al.* 2005.) In dogs, the linkage disequilibrium blocks, which are non-associated blocks of SNPs that are mostly inherited together, may extend to even up to mega bases, requiring only tens of thousands of SNPs to cover genome wide association studies completely, and therefore the research on the dogs' genetic variation is easier compared to humans. (Lindblad-Toh et *al.* 2005, Sutter *et al.* 2004.)

Genome association mapping is one of the most efficient methods to understand genetic and phenotypic variation. Instead of needing multiple related samples as in linkage mapping, association mapping requires only a population

based sample. (Sutter et *al.* 2004.) The genetic research for the desired mutation starts by identifying the approximate chromosomal region or haplotype shared with the trait. (Gray et *al.* 2010.) First the locus must be determined before reducing the area to the desired gene. (Rimbault & Ostrander 2012.) Next generation sequencing or fine mapping are efficient ways to determine the location of the desired mutation. (Rimbault & Ostrander 2012, Gray et *al.* 2010.)

By using the direct hybridization of the genomic targets to the sequences bound in the array, it is possible to identify up to 2.5 million SNPs in only one DNA sample, allowing large scale population studies. However, for the studies of the genome of the dog, usually only 172 000 SNPs are required to cover most of the traits. After the amplification and the hybridization of the DNA, the samples are placed on the chip for the hybridization, followed by single base extension and fluorescent staining. The mutated DNA will bind to the complementary, synthetically made sequence on the chip whereas the non-mutated sequence remains unbound. After the signal amplification, the sample is ready for scanning and analysis. The DNA microarray is an efficient method for DNA analyses and played had a major role in genomic research as well as in population genetics and copy number variant studies. (National Genome Research institute, UC Davis.)

3.2 Genetics of canine size

Canine size is a polygenic trait, which means that it is influenced by multiple genes. Almost half of the variance in weight can be explained by the polymorphisms in six loci found by fine mapping. These polymorphisms are located in or near the following genes, insulin-like growth factor (*IGF1*), insulin-like growth factor receptor (*IGF1R*), growth hormone receptors 1 and 2 (*GHR1* and *GHR2*), high mobility group AT hook 2 (*HMGA2*), stanniocalcin 2 (*STC2*), *SMAD* family member 2 (*SMAD2*). (Rimbault *et al.* 2013.)

The ancestral allele, defined as the allele found in the wolf, of these SNPs is associated with larger canine size whereas derived allele is mostly found in

smaller breeds. In fact, the smallest breeds have been found to carry at least four out of the seven derived markers and are significantly smaller than the ones with the heterozygous alleles. (Rimbault et *al.* 2012.). The seven markers mentioned above are believed to be responsible for up to 65 % of the size variance in dog breeds weighing less than 41 kg / 90 lb (Rimbault et *al.* 2013.)

In this study the findings of the six markers found by Rimbault et *al.* (2013) are repeated along with additional markers. In the following section, the importance of the six genetic markers found by Rimbault et *al.* (2013) for dog body size is introduced by examining the function of the marker genes at molecular level.

4 GENETIC VARIANTS PREVIOUSLY ASSOCIATED WITH CANINE SIZE

Seven markers *IGF1, IGF1R, GHR1, GHR2, HMGA2, SMAD2* and *STC2*, named according to their nearest genes, are able to explain 46-52.5 % of the size variation in dogs. Even though these genetic markers may not be in the protein coding area of the genome, they may still influence the gene function. (Rimbault et *al.* 2013.)

Wolf, the ancestor of the domestic dog, is homozygous to the ancestral allele in the size associated genes, and larger breeds commonly carry the same ancestral allele while smaller breeds carry the derived allele. Derived alleles of these genes predict smaller body size in canine and the more there are derived alleles of these markers, the smaller the dog is. (Rimbault et *al.* 2013).

4.1 Insulin-like growth factor (*IGF1*) and its receptor (*IGF1R*)

Insulin-like growth factor (*IGF1*) and its receptor (*IGF1R*) are found to have a major role in size prediction of dogs in several studies (Rimbault et *al.* 2013, Hoopes et *al.* 2012, Sutter et *al.* 2007). However, it seems that in order to reduce the body size, the derived *IGF1* allele needs be accompanied by at least one additional size reducing marker (Rimbault et *al.* 2013).

In molecular level, the insulin growth factor protein binds to the insulin growth factor receptor which is a tyrosine kinase signal transducer, promoting cell growth and cellular differentiation (Alberts et *al.* 2008). IGF1 is an extracellular signal protein that acts in the receptor tyrosine kinases. Both *IGF1* and *IGF1R* stimulate the survival and the growth rate in many cell types. (Alberts et *al.* 2008.)

The *IGF1R* marker is associated with reduced skeletal size in dogs and other mammals. In the derived allele of *IGF1R*, a highly conserved arginine amino acid is changed to histidine by a non-synonymous SNP. As a result of this func-

tional mutation, multiple hydrogen bonds are prevented from forming within the cysteine-rich domain in the ligand binding extracellular subunit of the receptor. (Rimbault & Ostrander 2012, Sutter et *al.* 2007.)

The missense mutation in the *IGF1R* marker results in reduced levels of the IGF1R protein, which possibly reduces the binding of IGF1 to IGF1R, hence decreasing the signals of *IGF1*. (Rimbault et *al.* 2013, Hoopes et *al.* 2012, Sutter et *al.* 2007.) Many canines has been found with size associated homozygous mutation in *IGF1R*, suggesting that despite the homozygous mutation in *IGF1R,* these individuals must have some *IGF1R* related function as they are alive. (Hoopes et *al.* 2012.)

4.2 Fibroblast growth factor (*FGF4*)

The breed-defining chondrodysplasia caused by the *fgf4* retrogene is a breed defining trait and the favorable phenotype known for its short legs, curved limbs and heavy bones. (Parker et *al.* 2009.) The *fgf4* retrogene causing breed defining chondrodysplasia found within several breeds, must not be mistaken to the short-limb causing chondrodysplasia, which is not a breed defining trait but a genetic illness. The chondrodysplasia that causes undesired dwarfism in some breeds, such as the Norwegian elkhound, is most likely caused by a mutation in the collagen-binding alpha subunit 10. (Kyöstilä et *al.* 2013.)

Breed-defining chondrodysplasia is caused by a 5 kb insertion containing a retrogene of the fibroblast growth factor (*fgf4*). The conserved retrogene contains the 3’ untranslated region and the polyadenylated tail (poly-A tail) of *FGF4*. However the 5’ regulatory sequences are not included in the *fgf4* retrogene. The retrogene may cause alteration in the *FGF4* function, causing excessive levels of the FGF4 protein, resulting malfunction in the FGF4 receptor timing during fetal development, leading to a premature closure of growth plates in the long bones, thus causing short-legged individuals. (Parker et *al.* 2009.)

The FGF4 protein belongs to the fibroblast growth factor group along with 24 similar proteins (FGF1-24). By inhibiting the differentiation of some precursor

cells or by acting as inductive signals in development, FGF4 proteins stimulate the proliferation of various cell types. (Alberts et *al.* 2008.) Like IGF1, the FGF4 is an extracellular signal protein that acts through the receptor tyrosine kinases (Alberts et *al.* 2008).

The *fgf4* retrogene has not been found in wolves, which most likely indicates that the duplication of the gene occurred soon after domestication. In addition, no individual has been found with the ancestral haplotype of *FGF4* and the *fgf4* insertion. (Parker et *al.* 2009.)

4.3 High mobility group- A2 (*HMGA2*)

The SNP in the 5' UTR of high mobility group- A2 (HMGA2) modifies the HMGA2 protein and so affects its function. Because HMGA works as a protein inducer, low levels of HMGA protein are visible in the phenotype (Rimbault et *al.* 2012, Hock et *al.* 2008), thus effecting the body size. In fact, *HMGA2* is associated also with height in humans but has not been shown to influence the body mass in adults (Weedon 2007).

High mobility group AT hook 2 protein is one of the four members of the high mobility group A (HMGA), which are chromatin binding proteins. All of the HMGA family group proteins have AT-hooks, which they use for the on-off binding to the AT stretches on the chromatin of the DNA and recruit additional components to the binding site. (Hock et *al.* 2008.)

4.4 Stanniocalcin (*STC2*)

The SNP marker named after its closest gene stanniocalcin (*STC2)* in this study is located 20 kb downstream from the protein coding region the of gene *STC2* but still it is believed to reduce the body size by altering the gene function (Rimbault et *al.* 2013).

Stanniocalcin-2 (STC2) is a widely expressed secreted polypeptide in mammalian tissue. It inhibits the activity of the proteolytic pregnancy-associated plasma protein-A (PAPP-A), which is a growth promoting metalloproteinase. STC2 inhibits PAPP-A by forming a disulfide bond with its Cys-120. By binding to the PAPP-A, STC2 prevents the normal growth by inhibiting the cleavage of the insulin-like growth factor protein (IGFBP)-4 and consequently the release of IGF. (Jepsen et *al.* 2014.)

4.5 Growth hormone receptor 1 and 2 (*GHR1* and *GHR2*)

The SNPs in the growth hormone receptors are protein altering thus affecting the body size (Rimbault et *al.* 2012). The growth hormone receptors belong to the family of the transmembrane proteins, and they are involved in various actions in the cell, by activating or inducing other signaling pathways. (Postel-Vinay & Kelly 1995.) The growth hormone receptors 1 and 2 (GHR1 and GHR2) are involved in the cell growth by activating the cell signaling pathway after the growth hormone binging (Alberts et *al.* 2008).

4.6 SMAD family member 2 (*SMAD2*)

The *SMAD2* marker is a deletion (9.9 kb) 15 kb away from the gene *SMAD2*. Despite its location far from the gene, it is believed to reduce the body size by affecting the transcription efficiency. (Rimbault et *al.* 2013.)

The SMAD family member 2 is involved in the regulation of the transcription of specific target proteins. It moves between the cell membrane and the nucleus according to the signal of the transforming growth factor- β hormone. The SMAD family members together with transforming growth factor- β hormones are responsible for the cell proliferation, cell specification and differentiation, extra cellular matrix production and cell death. In adults, the SMAD family members are responsible for the tissue repair and the immune regulation. (Alberts et *al.* 2008).

4.7 Additional markers

In addition to the markers described above, only recently eleven additional canine size associated markers have been identified (Hayward et *al.* 2016). As the locations of these markers have been published only lately, these markers have not been taken into consideration for this study. However, these markers are believed to be promising candidates for future studies of the size prediction in canine. In table 1, the locations of the newly discovered eleven size associated genetic markers are listed along with the information on the base of the ancestral and the mutated marker.

Table 1: The locations of the additional genetic markers.

Recently discovered size associated markers

CFA	loci	Ancestral	Derived
1	55983871	G	A
3	61986452	G	A
3	91103945	C	A
7	30243851	A	G
10	8183593	G	A
11	26929946	A	G
12	33733595	A	G
20	21479863	A	C
26	13224865	C	A
34	18559537	G	A
X	102212242	A	G

The reason for these markers affecting the body size is not yet fully understood for all variants but they are believed to explain nearly 90 % of the canine size variation in purebred dogs in combination with the previously described markers. (Hayward et *al.* 2016.)

5 MATERIALS AND METHODS

The total number of purebred canines used in this research is 16 372, which includes 77 wolves, 27 coyotes and 13 dingoes. The optimal number of dogs used in different analyses in this study was determined by the case, depending on the available successful genotyping results. However, for each section of the study as many canine as possible would be taken into consideration and a balance between the number of individuals and breeds was tried to be maintained. This section has the required information on the number of individuals in the research and the methods of the study in more detail.

5.1 Study Sample

Customer submitted samples for commercial DNA testing were made available for further analyses at Genoscoper. Most of the sample DNA was collected for the genotyping by using a buccal swab or in some rare cases, a blood sample.

The blood samples were taken by a vet once the identity of the canine was confirmed from the microchip of the dog. The buccal swab samples were taken by the dog owner or optionally by a vet. No separation on the sample quality according to whether the genotyping is done from a buccal sample or a blood sample was observed in this data.

5.2 Phenotype assignment

As done previously by Rimbault et *al.* (2012) and Boyko et *al.* (2010), to reduce the effects of the environment, the standard breed weight along with the standard breed height for each breed was determined by the kennel club standards or by other informative sources.

The standard breed weight (SBW) and the standard breed height (SBH) was assigned for each of the studied 283 breeds by using the breed standards of the

FCI and the AKC. If the breed did not have a specific SBW or SBH set by the FCI or the AKC, the complete dog breed book (Dennis-Bryan 2012) was used. In case of a new or a rare breed, the American rare breed association breed standards were used for the determining of the SBW and the SBH. For some popular breeds that have not yet been recognized by the kennel clubs, the breed club standards were used. The breeds were: Biewer Terrier (Biewer Club of America), Mi-ki (American Mi-Ki Registery association), Boerboel (Boerboel International), Kritikos Lagonikos (Kretahund), Bull Arab (Australian Bull Arab registery Inc.), English Toy Terrier (English Toy terrier club of the UK), Labradoodle (Australian Labradoodle association) and Russell Terrier (The Westminster Kennel Club). For the SBH and the SBW of the wolf, dingo and coyote, the Animal fact files: Mammals –book (Maailman nisäkkäitä) was used.

If different SBW or SBH was given separately for males and females, the average of these values was used by combining the lowest female value with the highest male value and dividing it by two. Breeds that had only the minimum or maximum height or weight in the breed standards, the value in question was used as the breed standard. These breeds were: Dogue de Bordeaux (minimum weight), Biewer Terrier (maximum height), and Toy Poodle AKC (maximum height), Irish Wolfhound (minimum weight), Old English Sheepdog (minimum height), Pekingese (maximum weight), Komondor (minimum height), Chinese Crested Dog (maximum size) Anatolian Shepherd Dog (minimum height).

If a breed had registry sub division according to size, mostly the medium or the standard size was used. These breeds were: Alaskan Klee Kai (Standard size), American Eskimo dog (Standard size), Manchester Terrier (Standard size), American hairless (Standard size), Rat terrier (Standard size), German Spitz: (Gross (large) size), Boston Terrier (Medium size), Xoloitzcuintle (standard size).

For clarity, only AKC standard size poodles were used for the analyses, excluding the FCI standard poodles as both poodles are genetically the same, even though their classification differs by size in each continent. All the breed weights and heights are untransformed, expressed in kilograms and centimeters.

5.3 Genotyping

The desired SNPs were analyzed with the commercially available Illumina microchip DNA array. In addition to the size associated SNPs, also some additional SNPs, which are not relevant to this study, were identified at the microarray genotyping.

5.4 Genetic markers

The size associated genetic markers, referred as the “six markers”, named according to their nearest genes, have been published previously in several of studies. One of the first dog body size associated genetic markers *IGF1,* originally discovered by Sutter et *al*. (2007) and the canine size reducing effect of *IGF1R* was first published by Hoopes et *al*. (2012), were again studied by Rimbault et *al.* (2013) along with four additional canine body size associated markers, named according to their nearest gene: *GHR1, GHR2, HMGA2* and *STC2*. These six markers were the main markers evaluated in this study.

The genotype of the *SMAD2* marker (deletion 24 kb down stream of *SMAD2*) which was one of the seven size associated markers discovered in the research of Rimbault et *al.* (2013), was unsuccessful with the chosen technology and for that reason it is left out of this research.

The locations of the six markers are stated on the table 2, along with the base of the ancestral allele and the mutated alleles. The positions are according to the CanFam 3.1 assembly version of the canine genome.

Table 2: The position of the main six markers analyzed in this study

Name	Ancestral	Derived	Position (CanFam 3.1)
IGF1	G	A	CFA15: 41,221,438
IGF1R	G	A	CFA3: 41,849,479
GHR1	G	A	CFA4: 67,040,898
GHR2	C	T	CFA4: 67,040,939
HMGA2	G	A	CFA10: 8,348, 804
STC2	T	A	CFA4: 39,182,836

In addition, the markers for the breed-defining chondrodysplasia were identified according to the previous findings by Parker et *al.* (2009).

5.5 Statistical analyses

SPSS statistics was used for the statistical analyses of the six markers. The Kolmogorov-Smirnoff test was used to determine whether the data is normally distributed in the six markers. Mann Whitney test was used to determine the difference between the two independent samples and the Kruskall-Wallis test on the three independent samples.

5.5.1 Marker analyses

Ten females and ten males were randomly selected from each breed for the marker analyses. Only breeds that had at least ten females and ten males were considered for the analyses. The total number of canine in the “Six markers” analyses is 3460, from 173 breeds. The random selection was done independently to ensure that most likely close family members were not selected and none of the breed was over-represented. If possible, individuals were randomly selected from different countries.

5.5.2 Frequencies

The frequencies of the six markers, chosen based on all successful genotyping data, were analyzed by using the genotypes of 15171 canine in 225 breeds. The same individuals are present in both the frequency and the fixed or nearly fixed –tables (table 4 and table 5). The frequencies were analyzed separately for each breed to find any fixed or nearly fixed alleles within each breed. The frequencies were determined according to the derived alleles, the more derived alleles, the higher is frequency.

The breeds were then divided into weight and height categories. Each height and weight was rounded to the nearest two (centimeter or kilogram) which formed one size category. The scale for the higher weight and height groups may be wider than the lower value groups because the number of individuals decreases in the larger size groups. For clarity, the more fixed the ancestral allele is, the darker blue the color is, and the more frequency for the derived allele, the darker red color is used in table 4 and table 5.

5.5.3 Breed defining chondrodysplasia

The methods for determining the breed-defining chondrodysplasia were evaluated according to the findings by Parker et al (2009). All 16 372 canine in total were tested for the breed defining chondrodysplasia genotype qualification

The logic used to define chondrodysplastic individuals and breeds was as follows:

1. The breed must have no *fgf4* signal on a markers designated to evaluate insertion locus
2. The breed must have no G/G in SNP3
3. The breed must have A/G in SNP1
4. The breed must have T/T in SNP2

The breeds that had at least 97 % frequency for all the requirements for chondrodysplasia were given the chondrodysplastic breed status. As the potential candidate breeds for the breed defining chondrodysplasia were considered breeds with the frequency of 46 % to 95 %. Breeds with a frequency less than 27 % were considered normal legged.

5.5.4 Combinations

For the analyses of the combinations of the six markers, 11009 dogs were taken into account based on their successful genotyping data in all of the markers in question.

The combinations were identified according to both homozygous and heterozygous alleles. In the homozygous combinations, the heterozygous alleles were considered as derived alleles, and in the heterozygous combination, the heterozygous alleles were considered their own, individual third allele type.

The combination frequencies were analyzed as well as the size distribution within each combination. In the combination tables, the derived alleles are colored gray, while the ancestral alleles are white. The order of the markers in the combination tables is from the top: *IGF1, IGF1R, GHR1, GHR2, HMGA2, and STC2.*

6 RESULTS

All weights and heights in this section are displayed in average standard breed weight and height (SBW, SBH). This section goes through all the results of the used analyses methods. First the box-plots are examined followed by the frequencies of the alleles. The combinations of the alleles are examined as heterozygous combinations as well as considering the states only homozygous. At the end of the section the results of the analysis of chondrodysolasia are introduced along with the newly discovered genetically short-legged breeds.

6.1 Association of individual markers with size

With the aim of analyzing the canine size distribution for each markers, the boxplots for each genetically different marker was tested. The following boxplots in figure 1 and figure 2 show the canine size distribution for each marker. The SBW or the SBH are shown on the y-axis while the alleles are on the x-axis.

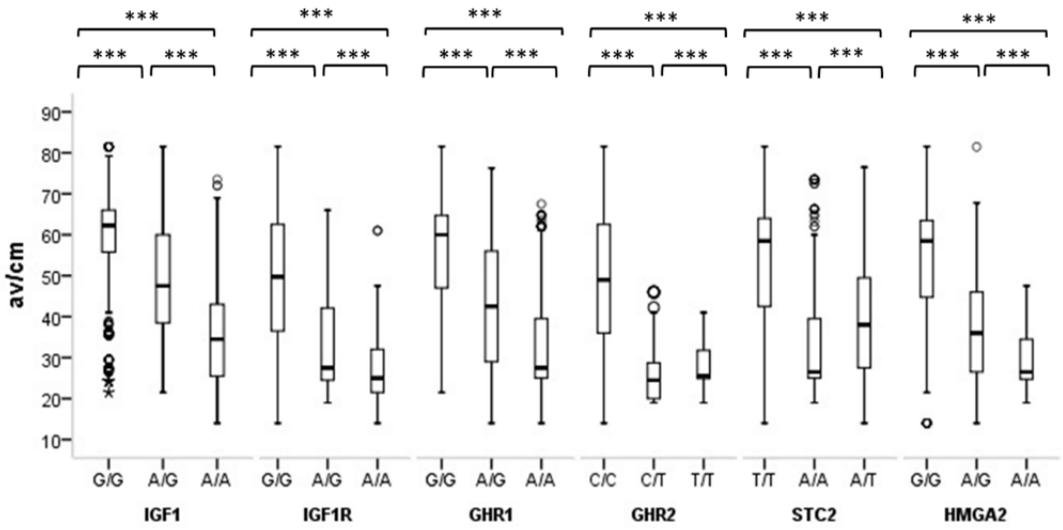

Figure 1: Canine standard breed height by genotype of the six markers based on 3460 individuals from 173 breeds. P-values statistically significant at the level of 0.001 were marked (***)

The derived alleles from each marker reduce the size of canine according to the Kolmogorov-Smirnoff test, the Mann-Whitney test and the Kruskall-Wallis test.

According to the Kruskall-Wallis test on the three independent samples the size of dog differ significantly depending on whether there is an ancestral, derived or heterozygous allele present in each marker.

Once noticed that the weight and height distribution differs in each marker group overall, the statistical test was continued by comparing three independent samples followed by the two independent sample comparison. When testing two independent samples by using the Mann Whitney, all but two pairs have a scientifically significant value.

As the null hypothesis is that there is no significant difference between the groups, the Mann Whitney results for the markers indicate a significant difference between the sizes of the dogs with different genotypes at the examined markers. Because there is no significant difference between homozygous derived allele and the heterozygous alleles at *IGF1R* and *GHR2,* most possibly even just one copy of the dominant, derived allele of *IGF1R* or *GHR2* reduces the weight of the dog.

The graphs (Figure 1 and Figure 2) for the weight and height shows statistical significance in all three statistical tests done with Mann Whitney, marked with three stars for the significance. The size distribution within the heights of each allele group is wider than with the weights. In fact, all the height groups were significantly different in all tests, as indicated with the P-values.

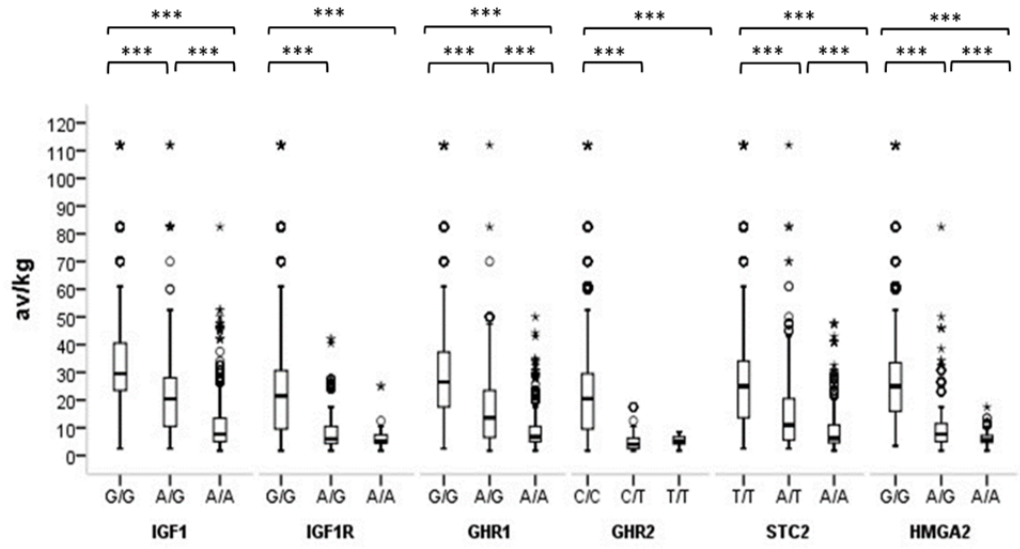

Figure 2: Canine standard breed weight by genotype of the six markers based on 3460 individuals from 173 breeds. P-values statistically significant at the level of 0.001 were marked (***)

The derived markers reduce the size of canine similarly in both height and weight. Very surprisingly *GHR2* heterozygous allele produces canine with the mean weight and height lower than the homozygous derived allele. The mean weight of the homozygous *GHR2* is 5.2 kg and mean height 24.5 while the mean weight for heterozygous *GHR2* is 4 kg and mean height 25.5 cm.

All the boxplots show that both weight and height are influenced by the markers. However, the plots show only the large picture by focusing on the means, and the effect of the other markers in the marker combination are not considered.

6.2 Fixed or nearly fixed alleles across breeds

Breeds that have a higher frequency of derived alleles tend to be smaller than the ones with more fixed ancestral alleles. However, some presence of the derived alleles can be found in all weight categories and also smaller dogs may have ancestral alleles despite their slightly smaller size, even though the ancestral alleles are associated with larger canine size.

The largest dog in weight with a completely fixed derived *IGF1* allele is the smallest of the Swiss mountain dogs, the Entlebucher Mountain Dog (24.5 kg, 46 cm). The fixed derived *IGF1* allele possibly explains the smallest size among the Swiss Mountain dogs. The frequency of derived *GHR1* allele in this breed is only 0.357 while all other markers were fixed to the ancestral alleles. German pincher (48 cm) is the tallest breed with fixed derived *IGF1* allele. Compared with the Entlebucher Mountain Dog, the German Pincher has fixed ancestral alleles only in *GHR2* and *HMGA2*, while in *STC* there is relatively high frequency of derived allele (0.85). Table 3 shows the largest breeds with each fixed derived allele.

Table 3: The sizes of the largest breeds with fixed derived alleles.

Fixed derived allele	**Largest breed**	**weight**	**tallest breed**	**height**
IGF1	Entlebucher Mountain Dog	24,5 kg	German Pincher	48 cm
HMGA2	Terrier Brazileiro	10 kg	Wirehaired Fox Terrier	38 cm
GHR1	Swedish Vallhund	14 kg	Norwegian Lundehund	34 cm
STC2	Griffon Bruxellois	4,8 kg	Phalene	28 cm

The tallest breed with a fixed derived *HMGA2* is the Wirehaired Fox Terrier (38 cm), and the heaviest breed is the Terrier Brazileiro (10 kg), indicating that derived *HMGA2* allele reduces the canine weight more than derived *IGF1* allele. The tallest breed with a fixed derived *GHR1* allele is Norwegian Lundehund (34 cm), and the heaviest Swedish Vallhund (14 kg), and for the derived *STC* allele the tallest is the Phalene (28 cm) and the heaviest Griffon Bruxellois (4.8 kg). By examining the fixed alleles it may be observed that the largest breeds with fixed derived alleles are in fact small compared to the size of wolf.

Out of the 225 breeds in which allele frequency could be calculated in, up to 190 breeds are fixed for the ancestral *GHR2* allele, while no breed is fixed for the derived *GHR2* allele. For *IGF1R*, 133 breeds are fixed to the ancestral allele and similarly to *GHR2*, no breed is fixed for the derived allele in *IGF1R*. Despite the fact that derived I*GF1* allele is the most common derived allele, only 34 are breeds are found completely fixed to it, while 28 breeds are fixed for the ancestral *IGF1* allele. Figure 3 shows the number of breeds with fixed alleles. Derived *GHR1, STC2* and *HMGA2* alleles, were found in 6, 2 and 9 breeds respectively, while ancestral alleles were found in 45, 54 and 103 breeds, respectively.

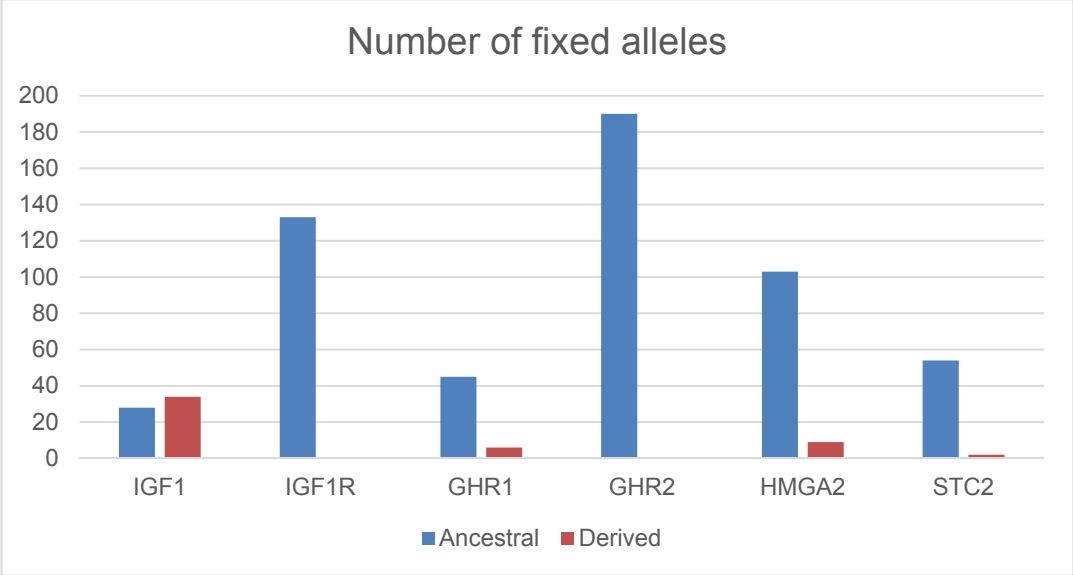

Figure 3: Number of alleles fixed for a breed. The number of fixed ancestral alleles out of 225 breeds are in blue and the fixed derived alleles in red.

Rimbault et *al.* (2013) found only a few dogs that were heterozygous to *HMGA2, IGF1R and GHR2* markers. In fact, *HMGA2* marker had only 15 heterozygous genotypes, whereas homozygous derived genotypes alleles were observed 87. In the data of over 16 000 individuals used in this study, the derived allele of *HMGA2* was found 2700 times and heterozygous allele 1171 times, giving the same ratio of derived and heterozygous alleles as found previously.

The average size of a dog with a fixed derived *HMGA2* is 6.4 kg and 30 cm, average size with canine with a fixed *GHR*1 is 6.4 kg and 25.9 cm. The average size of a dog with derived *IGF1* allele is 6.5 kg and 31.4 cm. All breeds with any fixed derived allele, with the exception of *STC2*, have the mean weight of approximately 6.5 kg, while there is some height difference between the means of the heights.

The only two breeds with a fixed derived *STC2* allele are Griffon Bruxellois (4.75 kg, 26 cm) and Phalene (3.5 kg, 28 cm), making the average weight of the breeds with a fixed *STC2* allele 4.1 kg and 27 cm. Both of the breeds are also fixed to the derived *IGF1.* In addition to derived *IGF1*, Griffon Bruxellois is fixed to the derived *GHR1* and nearly fixed to the derived *GHR2* (frequency above 0.97). Phalene is only fixed to the ancestral *GHR2 and IGF1*. However, Griffon Bruxellois is fixed to the ancestral *IGF1R*, which might explain the similar body size in these two breeds.

Some interesting specific observation was that the only two ancestral alleles that are fixed in larger breeds alone are *IGF1R* and *GHR2*. Therefore these two markers seem to have the strongest impact on reducing the body size. American Akita (41 kg, 66cm) is the largest dog (in height) having some derived *IGF1R* alleles within the breed (frequency 0.022). The largest breed in which the derived *GHR2* allele is found (frequency 0.033), is Xoloitzcuintle (15 kg, 53 cm), indicating that *GHR2* has a strong effect in reducing canine size. In fact Xoloitzcuintle is also the largest breed in which all the derived alleles are found.

Similarly to the previous findings, 92% of the Rottweilers (46 kg, 62 cm) were carrying a derived *IGF1* allele (Rimbault et *al.* 2013). However, *IGF1R, GHR1* and *GHR2* markers in Rottweilers are fixed to ancestral allele and *STC2* and *HMGA2* have relatively small frequencies (0.005 and 0.027 respectively) for derived alleles. Among the larger breeds, Dogue de Bordeaux (48 kg, 63 cm) has relatively high frequency of the derived *IGF1* allele; nearly half of the individuals are carrying the derived *IGF1*. The derived *IGF1* allele is however the only derived allele Dogue de Bordeaux is carrying and none of the members of that breed carry any other derived allele.

The examining of the fixed alleles gives valuable information on the genetic markers determining the sizes of different breeds. Derived alleles are more commonly found fixed to the smallest breeds. If a derived allele of a marker is only rarely found fixed to a breed, the derived allele of the marker seems to result in smaller sized dogs. If the derived allele is found fixed for many breeds, the size reducing effect of the allele can be considered moderate. The fixed derived alleles indeed seem to reduce the overall size of the breed while the fixed or nearly fixed ancestral alleles are found in nearly all size categories but mostly within the larger breeds.

6.3 Frequencies of derived alleles

For the frequency tables the individual genetic data along with the size information of each breed was divided into weight or height groups of each 2 centimeter or 2 kilograms. The alleles and their frequencies were calculated separately from each other. When examining the frequencies of the markers separately according to the weight and height of the canine, the frequency of the derived alleles is higher within the smaller breeds compared to the larger size ranges. In addition, no height or weight range has any fixed derived alleles to that category, but most of the ancestral alleles are fixed within the categories of the larger breeds.

Table 4 presents the allele frequencies of the height groups and table 5 the frequencies for the weight groups. The tables show only the derived alleles. The darker the red is, the higher the frequency of the derived allele is. The blue color indicates a small frequency of the derived allele and the darkest blue means there is no or almost none derived alleles in that group. In the tables, it may be seen that the frequency of colors gets stronger towards the extremes of the size ranges, while in the middle sized groups, the frequencies are mostly not fixed and thus in light colors.

Table 4: Allele frequency of the six markers in each height group.

cm	*IGF1*	*IGF1R*	*GHR1*	*GHR2*	*STC2*	*HMGA2*	Count	Breeds
75+	0,01	0,00	0,01	0,00	0,02	0,01	302	6
72-74	0,06	0,00	0,01	0,00	0,13	0,00	450	8
68-70	0,02	0,00	0,04	0,00	0,01	0,00	518	9
66	0,16	0,00	0,04	0,00	0,07	0,00	550	10
64	0,09	0,00	0,08	0,00	0,05	0,01	1211	15
62	0,22	0,04	0,11	0,00	0,05	0,01	844	16
60	0,14	0,00	0,11	0,00	0,15	0,00	786	11
58	0,26	0,02	0,15	0,00	0,08	0,01	367	9
56	0,19	0,00	0,15	0,00	0,06	0,02	1351	9
54	0,52	0,01	0,26	0,00	0,10	0,02	141	4
52	0,47	0,00	0,14	0,00	0,48	0,03	270	3
50	0,49	0,00	0,36	0,00	0,18	0,01	766	8
48	0,54	0,04	0,24	0,00	0,23	0,03	613	11
46	0,79	0,02	0,20	0,01	0,18	0,06	787	10
44	0,74	0,20	0,27	0,00	0,21	0,05	1053	7
42	0,85	0,15	0,65	0,07	0,20	0,08	484	8
40	0,87	0,06	0,44	0,00	0,17	0,26	334	4
38	0,84	0,01	0,21	0,00	0,21	0,61	417	7
36	0,80	0,07	0,42	0,00	0,29	0,29	696	10
34	0,98	0,07	0,51	0,03	0,33	0,82	456	5
32	0,99	0,13	0,90	0,32	0,51	0,29	357	4
30	0,79	0,02	0,42	0,00	0,30	0,46	350	7
28	0,75	0,10	0,73	0,04	0,50	0,66	731	11
26	0,94	0,12	0,81	0,12	0,66	0,86	972	15
24	0,81	0,22	0,89	0,24	0,55	0,80	240	5
22	0,86	0,47	0,71	0,07	0,57	0,70	258	6
<20	0,96	0,45	0,90	0,26	0,72	0,76	413	7
							15717	225

The chondrodysplastic breeds with short legs are also part of this table and it may affect the results by showing the larger body sized dogs in smaller height groups due to the shorter legs.

The table of the height categories shows that the highest frequencies of the derived alleles are not in the smallest size groups but somewhere around the

height of about 30 centimeters. The derived *IGF1* can be found in all the height categories however, the highest frequency of derived *IGF1* allele is found in the height category of 32 cm with frequency above 0.99. The derived *IGF1R* allele has the highest frequency (0.45) at 22 cm, the derived *GHR1* allele at 32 cm (0.90), the derived *GHR2* allele at 32 cm (frequency 0.32), the derived *STC2* allele at below 20 cm (frequency 0.72) and the derived *HMGA2* allele at 26 cm (frequency 0.86). Derived *HMGA* allele frequency is below 0.01 for the breeds that are at least 58 cm in height. The frequencies for derived *HMGA* allele are above 0.1 for the breeds with the average height less than 40 cm.

The frequencies for each height category group for each 2 cm shows that there is only ancestral *GHR2* alleles present until canine of 54 cm in height indicating that *GHR2* is the least frequent derived allele found. In fact, even within the smallest breeds, the maximum frequencies of ancestral *GHR2* are below 0.3. Rare derived *IGF1R* allele can also be found within the smaller breeds with frequencies below 0.5. This possibly indicates that these two rare alleles have a strong size reducing effect.

The frequency of the derived *IGF1* allele is almost or above 0.5 in all height categories until 56 cm. In breeds with average height of above 56 cm, the frequency for derived *IGF1* allele is below 0.3 and in larger breeds (at least 68 cm) the frequency of the derived *IGF1* allele is below 0.1. The frequencies for the derived alleles of *IGF1* in the weight categories are more complex. The derived *IGF1* allele frequencies are found to be below 0.1 from the canine weight of 32 kg to 38 kg and above 50 kg excluding the 65 kg to 75 kg weight category as shown in table 5. The frequency of the derived allele of *IGF1* is surprisingly more popular within the weight categories from 40 to 50 kg (frequency of above 0.1 and up to 0.35) and 65 to 75 kg (frequency 0.1). The frequency of the derived *IGF1* allele is above 0.90 on all weight categories below 6 kg. Table 5 shows the allele frequencies in all weight categories. The higher frequency for the derived allele is marked in darker red while the higher frequency for the ancestral alleles are marker in darker blue.

Table 5: Allele frequency of the six markers in each weight group.

kg	*IGF1*	*IGF1R*	*GHR1*	*GHR2*	*STC2*	*HMGA2*	Count	Breed
70+	0,12	0,00	0,01	0,00	0,03	0,01	453	4
62-68	0,01	0,00	0,01	0,00	0,01	0,01	287	5
52-60	0,11	0,00	0,03	0,00	0,00	0,00	140	5
48-50	0,10	0,00	0,05	0,00	0,06	0,00	289	7
44-46	0,28	0,00	0,13	0,00	0,05	0,01	401	6
40-42	0,11	0,01	0,04	0,00	0,22	0,00	267	5
36-38	0,02	0,00	0,04	0,00	0,03	0,00	419	6
32-34	0,08	0,00	0,12	0,00	0,04	0,01	755	8
30	0,23	0,00	0,15	0,00	0,07	0,01	1482	12
28	0,35	0,01	0,10	0,00	0,14	0,00	838	13
26	0,26	0,09	0,05	0,00	0,05	0,01	652	13
24	0,30	0,00	0,23	0,00	0,14	0,01	1098	12
22	0,16	0,00	0,16	0,00	0,08	0,00	511	9
20	0,62	0,00	0,18	0,00	0,16	0,00	731	11
18	0,80	0,05	0,31	0,02	0,43	0,07	255	7
16	0,64	0,02	0,46	0,00	0,21	0,01	626	11
14	0,71	0,17	0,30	0,00	0,22	0,03	1214	9
12	0,80	0,07	0,60	0,00	0,27	0,24	729	9
10	0,84	0,13	0,49	0,01	0,30	0,31	1140	18
8	0,77	0,09	0,64	0,05	0,34	0,81	560	11
6	0,94	0,10	0,73	0,12	0,54	0,81	1521	23
4	0,96	0,27	0,79	0,15	0,64	0,72	1002	16
2	0,97	0,28	0,87	0,38	0,83	0,94	347	5
						Total:	**15717**	**225**

The table of the frequencies of the alleles with the weight groups is very similar with the result as the height groups. However, because the chondrodysplasia does not have a strong impact on the weight, the table of the weights is more linear than the table of heights. In the table 5, it can be seen that the smaller the body size is, the higher is the frequency of the derived alleles in that group.

The weight range group that has the most fixed ancestral alleles is the group that weighs between 52-60 kg, which is in fact the size range of the wolf. This group in fact is not the largest weight group but the third largest. Similar result can be found in table of the height groups. The height group that has the highest frequency of ancestral alleles (no derived alleles at *IGF1R, GHR2, and HMGA2*) are in height 72 cm to 74 cm, which again is not the tallest height group. This may verify the hypothesis by Rimbault et *al.* (2013) that the larger breeds must have additional markers making them larger than their ancestors. The smallest weight group that carries the highest number of fixed ancestral alleles (which include ancestral alleles in *IGF1R*, *GHR2, and HMGA2)* is weighing 22 kg.

Canine weight seems to correlate with the derived allele frequencies more obviously than height. All the highest frequencies of the derived alleles are within the smallest weight range while there is more variation in the frequencies of the derived alleles within the height categories. The rarest alleles again are *GHR2, HMGA2* and *IGF1R*, which are found only within the tiny breeds. These three alleles seem to have the strongest size reducing effect.

6.4 Association of combinations with size

Within 11009 dogs of which all six markers successfully genotyped, 265 different combinations were found when considering the heterozygous state separately from the homozygous state. If the heterozygous state was combined with derived allele homozygosity, only 41 combinations exist, which in fact is two more than 39 combinations observed in the previous research by Rimbault et *al.* (2012). However, even though all the heterozygous states reduce the dog size compared to the ancestral states, it seems likely that the heterozygous states affect the phenotype slightly differently than homozygous derived states.

In the data of 11009 individuals, within the least frequent combinations, having only one carrier each, the median weight is 8.5 kg and height 28 cm. Within the combinations that have less than ten carriers each (total number of individuals:

508), the median for weight is 6 kg and the height 27.5 cm. This indicated that the rarest combinations can be found in the smallest individuals.

The marker combination with the largest amount of carriers (2852 individuals) is the full row of only ancestral alleles (GGGCGT). The weight distribution for this group is from the 10 kg of the Cirneco dell´etna up to over 120 kg Tibetan Mastiff. The shortest dog with this combination is Welsh Corgi Cardigan, with the average height of 29.5 cm, the breed however has the breed defining chondrodysplasia, thus shorter legs.

Surprisingly, one French bull dog (11kg, 29.5 cm) shows presence of the ancestral combinations and has the same height as the Welsh Corgi Cardigan. The reason for the unusual sized breed with this combination is yet unknown. The French Bulldogs are not fixed to any combination, but usually it has been found with at least two heterozygous derived markers, mostly at least one of them being *IGF1*. The French Bulldog and its larger sized ancestor, the English Bulldog (The French Bulldog club of America), are relatively short breeds, yet still they do not have the breed defining chondrodysplasia, which causes short legs. Most likely, there are some other unknown genetic markers in the bull dogs causing the short, yet muscular body. It may also be possible that this kind of exception would explain the individual size variation within the breeds. Figure 4 shows all the combinations found in the French Bulldog and the English Bulldog.

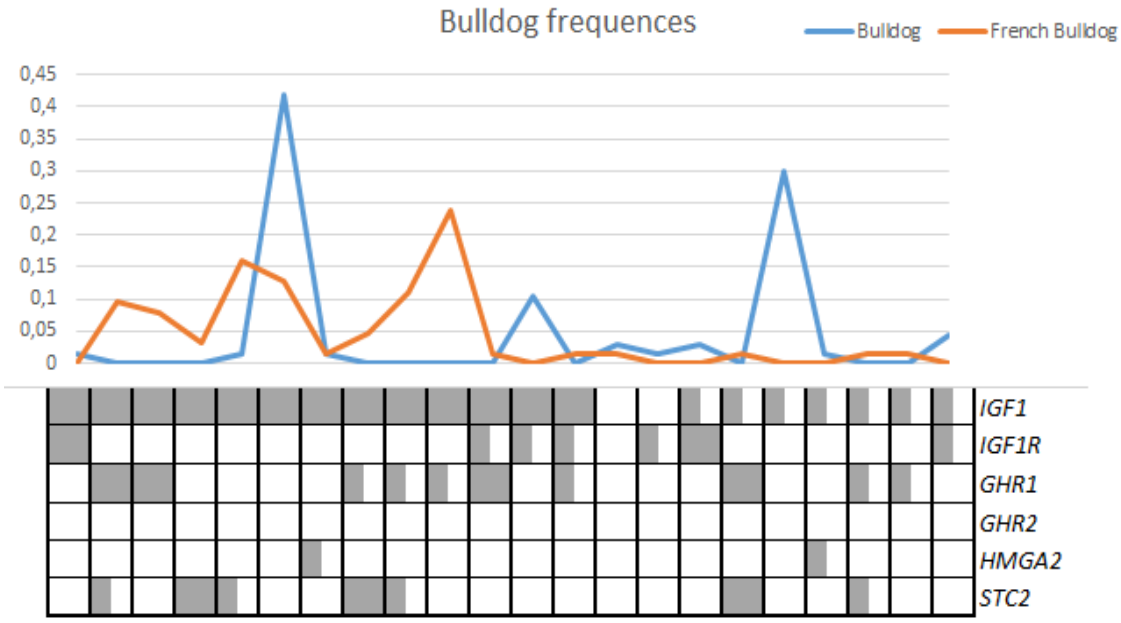

Figure 4: The combinations of the French Bulldog and the (English) Bulldog

The frequencies of the Bulldog combinations shows the combinations on the x-axis and the frequency on the y-axis. Blue line shows the Bulldog and the orange line shows the French bull dog. The derived alleles are marked as grey, ancestral alleles in white. All heterozygous alleles are marked with a half white half grey boxes.

The graph of the bulldog frequencies shows the alleles in the marker combinations. Mostly, the French bull dog is found with derived *IGF1* allele and with at least one other derived heterozygous or homozygous allele in any of the other markers, excluding derived *GHR2* allele. The English bulldog on the other hand is found mostly with only a derived heterozygous or homozygous *IGF1* with all other markers ancestral. For these combinations the size scale is wide, allowing both medium to large size dogs.

Within all the dog combinations, the three most common combinations are the all ancestral GGGCGT and the all ancestral combination with only *IGF1* marker either heterozygous (marked as R for RGGCGT by UIPAC nomenclature) or homozygous derived allele (marked as A for AGGCGT by UIPAC nomenclature). Even though the distribution of these combinations is one of the widest, in

fact the most popular ancestral combination cannot be found in breeds weighing less than 10 kg, while the combination with heterozygous *IGF1* marker may be found in breeds weighing at least 8.5 kg. The homozygous derived allele of *IGF1* with all other alleles ancestral may be found in breeds with the minimum average weight of 6 kg. The maximum weight for all of these combinations is over 100 kg, generally with this data meaning that they have been found in Tibetan Mastiffs. The three most common combinations are shown in Figure 5.

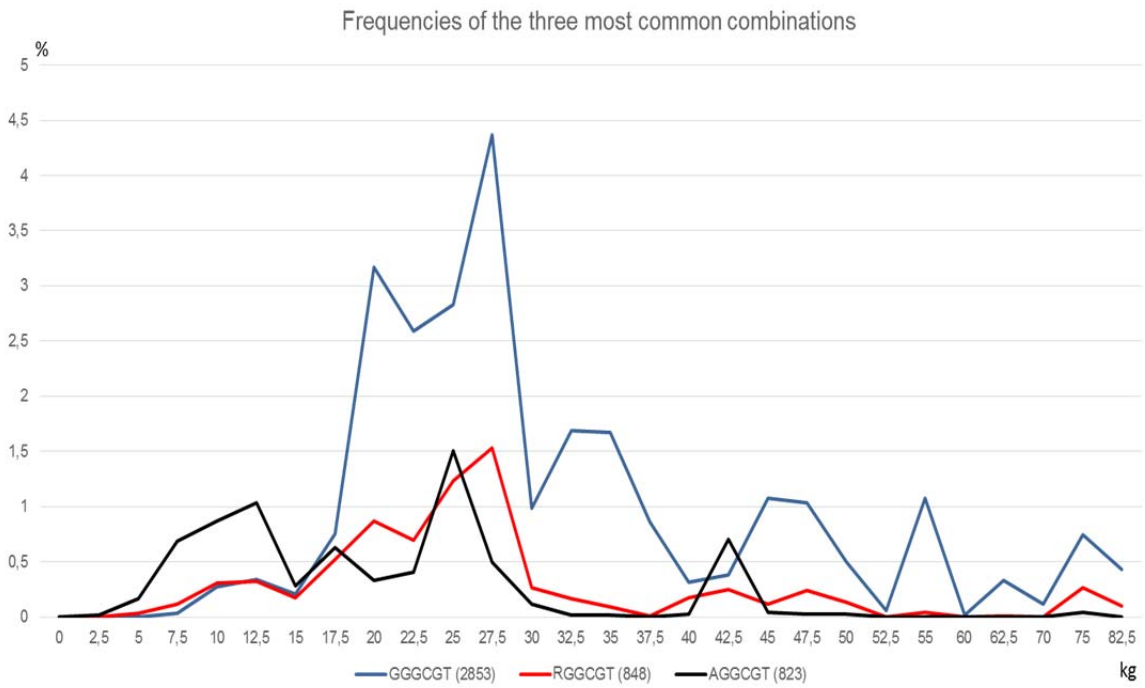

Figure 5: The frequencies of the weights of the three most common combinations

The interesting aspect of the graphs in Figure 5 is the peak of the combination frequency. All three combinations peak at about 25 kg, however the variation of the size of dogs in the all ancestral combination is wider compared to the other combinations. The peak for the ancestral combination with derived *IGF1* allele at about 46 kg is the Rottweiler, found with nearly fixed derived *IGF1* allele. 7.8 % of all canines carry the AGGCGT combination of which the most (11.4 %) is weighing 28 kg, out of which up to 99 % is American Staffordshire terriers and only one percent of Belgian Malinois.

One quarter of the 11009 canine has all ancestral alleles in the six markers. Within the largest group, no weight group stands out with a high frequency. The weights are equally distributed, with a slight peak at 29.5 kg with almost 7 % of the dogs with this combination belonging to this weight group. Out of the total canine amount nearly eight percent of the canines belong to the ancestral group with heterozygous derived *IGF1*. The most common weight group with this marker combination is about 30 kg, with only two closely related breeds in this group, the Labrador retriever (77%) and the Golden retriever (23 %).

Only two dogs are found with all ancestral alleles but derived *IGF1R* allele, one Lagotto Romagnolo (breed average 13.5 kg, 44.5 cm) and one shorthaired Vizsla (breed average 25 kg, 61 cm). The 119 canine that have all five ancestral alleles and derived *STC2* allele, have a weight range from 10 kg to 47.5 kg. Seven canine, weighing between 8.0 kg and 11.5 kg, were found with a combination of all five ancestral alleles and a derived *HMGA2* allele. Table 6 shows the SBW and the SBH of the individuals with the derived *HMGA2* allele in all ancestral combination.

Table 6: *HMGA2* marker as the only derived marker in a combination

Breed	Av/kg	Av/cm	*IGF1*	*IGF1R*	*GHR1*	*GHR2*	*STC2*	*HMGA2*
Cirneco dell' Etna	10,0	47,0	G/G	G/G	G/G	C/C	T/T	A/A
Fox Terrier - Wire	8,0	38,0	G/G	G/G	G/G	C/C	T/T	A/A
Shetland Sheepdog	11,5	36,3	G/G	G/G	G/G	C/C	T/T	A/A
Shetland Sheepdog	11,5	36,3	G/G	G/G	G/G	C/C	T/T	A/A
Shiba	9,2	38,0	G/G	G/G	G/G	C/C	T/T	A/A
Shiba	9,2	38,0	G/G	G/G	G/G	C/C	T/T	A/A
Welsh Terrier	9,3	38,5	G/G	G/G	G/G	C/C	T/T	A/A
total mean	9,8	38,9						

Based on the data of seven canine found with the only derived allele in *HMGA2*, the weight may be reliably predicted in about 10 ± 2kg.

No canine was found with the combination that had all ancestral markers but derived *GHR2* allele. The *GHR2* seems to be a special case. When *GHR2* allele is derived (homozygous T/T or heterozygous C/T), then *GHR1* is also de-

rived (homozygous A/A or heterozygous A/G). If *GHR2* allele is the homozygous derived allele (T/T) then *GHR1* is also homozygous derived allele (A/A).

Whether the marker in *GHR2* is derived or ancestral, has however no affect on the markers in *IGF1, IGF1R, HMGA2* or *STC2*. Derived *GHR1* allele allows both derived and ancestral allele in *GHR2,* but the ancestral *GHR1* is possible only if *GHR2* is also ancestral. Similar results were found by Rimbault et *al.* (2013), as they could not find any canine with the ancestral *GHR*1 and derived *GHR2* (G/G and T/T) combination. These two alleles are nonsynonymous, however, one possibility for the nonexistent GT haplotype is that *GHR2* variant was developed later than *GHR1* and only among the derived *GHR1* carriers. (Rimbault et *al.* 2013.)

Norwegian Lundehund (6.5 kg, 34 cm) is the only breed found with all alleles fixed to either derived or ancestral allele. The fixed derived alleles are *IGF1, GHR1* and *HMGA2,* while the ancestral alleles are *GHR2, STC2* and *IGF1R*. The reason for the fixed combination within the Norwegian Lundehund is due to the severe population bottlenecks it has gone through in recent times. (Institute of canine biology).

Similar combination to the Norwegian Lundehund was however found also in 30 additional breeds, yet not fixed to any other breed. In all canine, this combination is shared with the weight range of 2.5 kg to 12.5 kg, and 14 cm to 40 cm. In fact, only the combination of the Norwegian Lundehund and the all ancestral combination are the only combinations found fixed to any breeds.

The combination of all six ancestral markers is the most common combination (2852 individuals) of all data. The breeds that are fixed to all ancestral alleles are Pyrenian Mastiff (76 cm, 62 kg), Tamaskan dog (70 cm, 38 kg), Gray wolf (68 cm, 45 kg), Beauceron (66 cm, 34 kg), Saluki (64 cm, 23 kg), Akita Inu (64 cm, 40 kg and English Fox Hound (62 cm, 30 kg). The average heights of these breeds are very near to the height of the wolf, the ancestor of the domestic dog. These breeds are also mostly considered as the wolf-like or ancestral-like breeds. The smallest breed having all but one ancestral allele is the Harrier (52

cm, 24 kg). Yet still, the frequency of the derived *IGF1* allele in Harrier is only 0.21.

In the data of 11009 canine, only ten individuals in total were found with all derived alleles. These include seven Affenpinschers, two Malteses and one Chihuahua, with the average weight and height of 5 kg and 25 cm, 3.5 kg and 25 cm, 1.8 kg and 19 cm, respectively. In fact, none of these three breeds have any fixed alleles, although the derived *IGF1* allele in Chihuahua is nearly fixed (frequency above 0.99). An interesting aspect for the future would be the studies on whether these individuals predict smaller body size within the breed.

The frequency of ancestral alleles in the wolf was 1.00 in the 12 wolves genotyped for the six size associated markers as well as the derived allele frequency in the 13 tested dingoes. Out of 27 tested coyotes two were heterozygous for the derived *IGF1* allele and one individual was heterozygous for *GHR1* marker. In the previous studies by Rimbault et *al.* (2013), Coyote (2 individuals) and dingo (2 individuals) were found with fixed ancestral alleles in the six markers, while the wolf (26 individuals) was found with the frequency of 0.98 for the ancestral *IGF1* allele and ancestral *STC2* allele, other frequencies being 1.00. The differences are likely due to the different sized data.

The most frequent combinations are most commonly found within the larger sized dogs. The graphs in the figure 6 and figure 7 show both weight and height of the combinations found in at least 100 individuals. These combinations include 7889 canine from 247 breeds. The allele combinations in graphs are sort by the size from small in left to larger sized canine in right. Derived alleles are colored in grey, ancestral alleles in white.

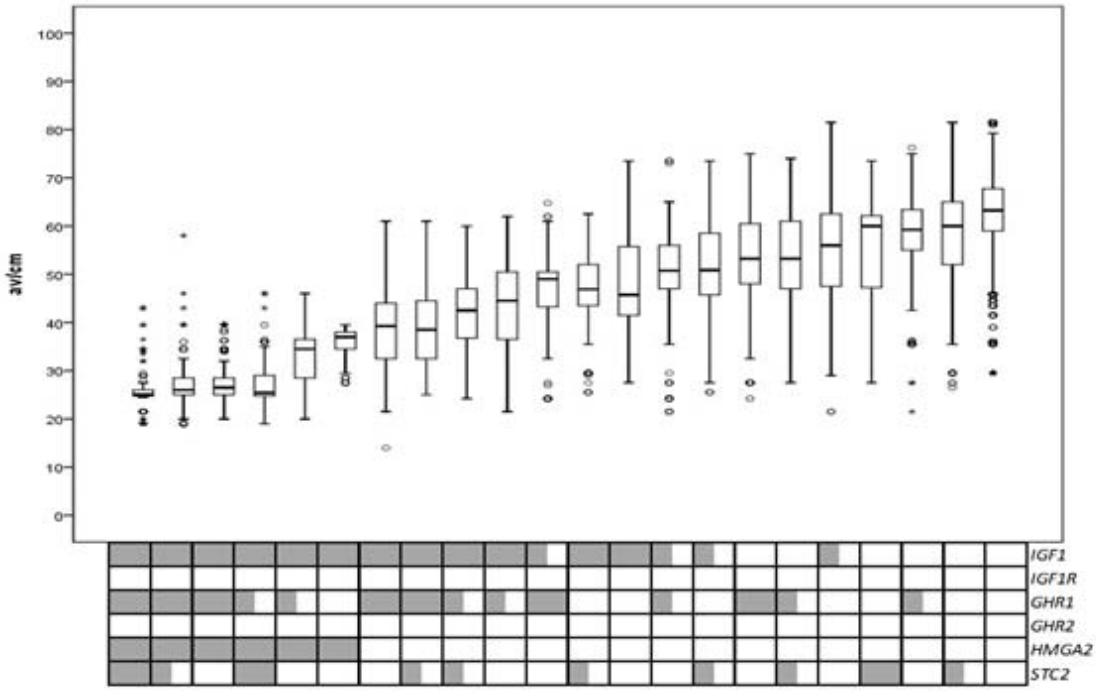

Figure 6: The height distribution of the combinations found in at least 100 individuals. Derived markers marked in grey, ancestral markers in white.

The breed defining chondrodysplasia, causing short legs in certain breeds, may affect the out-come of the height distribution. The heights of the different combinations seem to follow a liner and straight forward trend, more derived alleles in a combination results in a smaller body size.

According to the box-plots, even only one derived allele in an ancestral combination reduces the size of a dog. The more there is derived alleles, the more the size is reduced. In fact the number of the derived alleles seems to be the most important factor when determining size but some markers still have stronger size reducing effect than others.

The derived *HMGA2* allele accompanied by the derived *IGF1* allele seems to have the most significant effect in reducing size, compared to any other marker pair found in the most frequent combinations. In addition, the size distribution of dogs with a combination that has the derived *HMGA2* allele accompanied by the derived *IGF1* allele is very narrow, possibly allowing a fine scale size prediction.

The weight can be predicted similarly to the height with the same combination, and in the chart the same combination results in similar, balanced sized dogs in both height and weight as seen in the previous figure 6 and the following figure 7 which shows the weight distribution of the most common combinations.

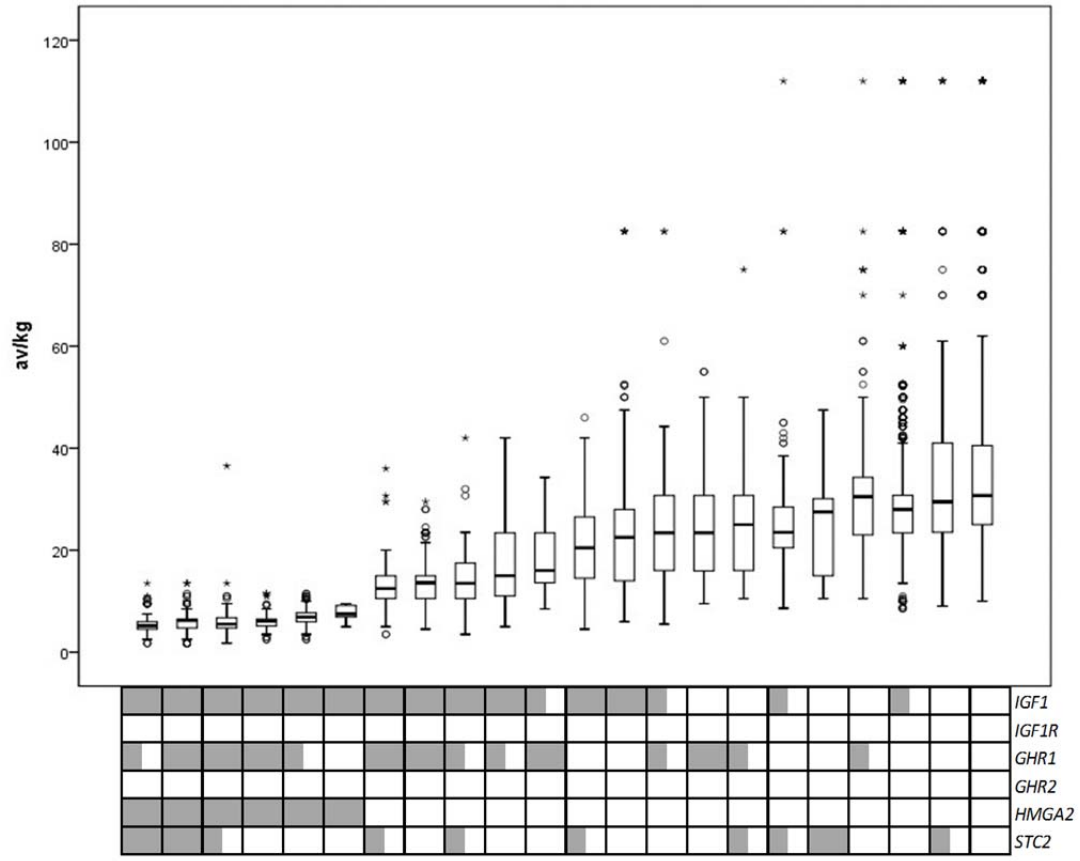

Figure 7: The weight distribution of the combinations found in at least 100 individuals. Derived markers marked in grey, ancestral markers in white.

The overall chart of sizes seem to be produced by similar order of combinations with only a few exceptions and indeed most of the statements made for the weight distributions seem to apply also to the height. All alleles seem to reduce size similarly in height and weight and the number of the alleles seems to determine both height and weight of the dog more importantly than the identity of the individual allele. Some specific observations on the reducing effect of the alleles or their combination can be however made.

As displayed in the graphs, the more there are derived alleles in the combinations, the smaller the SBW and SBH of the dog is. Within the most common combinations, also found within the largest breeds, the rare derived *IGF1R* allele and derived *GHR2* allele cannot be found. This indicated that these alleles are present only in the small sized dogs.

The weight of the canine is reduced with homozygous derived *STC2* allele, but the reduction is only weak with heterozygous *STC2*. However, when heterozygous *STC2* marker is accompanied by either heterozygous or homozygous derived *IGF1*, the size of canine is considerably reduced, especially in comparison to the ancestral combination with derived *IGF1* allele the only derived marker. This possibly indicates that derived *STC2* allele must be accompanied by another derived marker to reduce size. However, when there are more than three derived markers in addition to the derived copy of *STC2,* the reducing effect of *STC2* in the weight is almost negligible. Possibly it indicates a weak reducing effect of the derived *STC2* copy, which gets hidden by other markers when accompanied by more than three markers. However, this does not apply to the height as seen in the figure 6.

The rare combinations were examined with the marker combinations that had 12 to 28 carriers shown in figure 8 and figure 9. The combinations that had at least 10 carriers were considered to have some scientific variation, however only the combinations that had less than 30 carriers were considered rare.

Within the combinations of the 12 to 28 individuals in each combination shown in figure 8 and figure 9, only 780 individuals within 124 breeds were found. Most of the rare combinations are found in the breeds with average weight below 15 kg, while there is more variation in height.

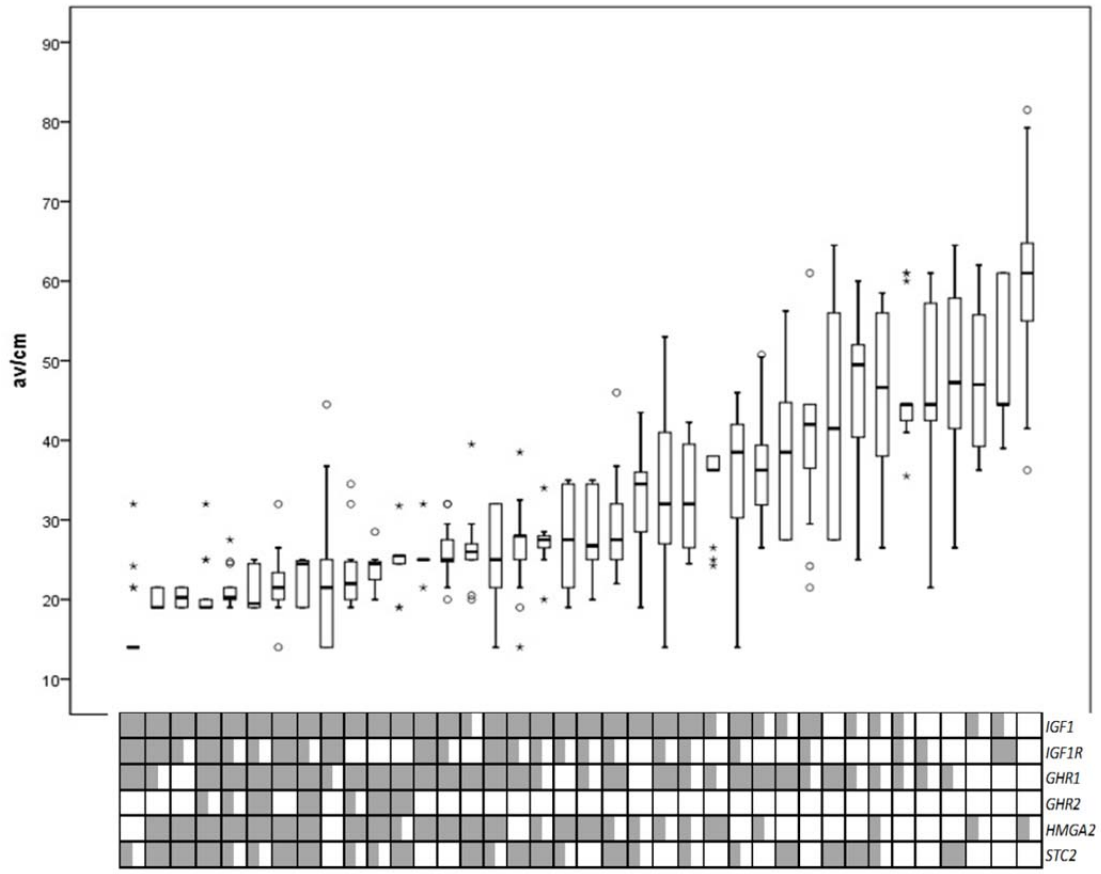

Figure 8: The height of the dogs with the combinations found in 12-28 carriers Derived markers marked in grey, ancestral markers in white.

Similarly to the most common combinations, also within the rarest combinations it may be seen that the more there is derived alleles in the combination, the more the size is reduced. The chondrodysplasia status may affect the outcome of the box-plot of the height of the dog with the rare combinations, especially within the smallest heights.

As seen in figure 8, the height of the dog is not strongly reduced with a heterozygous derived allele, however, if accompanied by other markers, the heterozygous alleles may reduce the size of dog if there are other derived alleles present in the combination.

According to the graphs in figure 8 and figure 9, the derived *IGF1R* allele and the derived *GHR2* allele are not within the most common combinations but only within the combination which show less than 30 individuals in each combination,

which again indicates that the derived *IGF1R* allele and derived GHR2 allele are rarely found markers.

Figure 9 shows the weight distribution of the rare combinations. As seen in the figure, the same combination results in a similar, balanced body size in both height and weight.

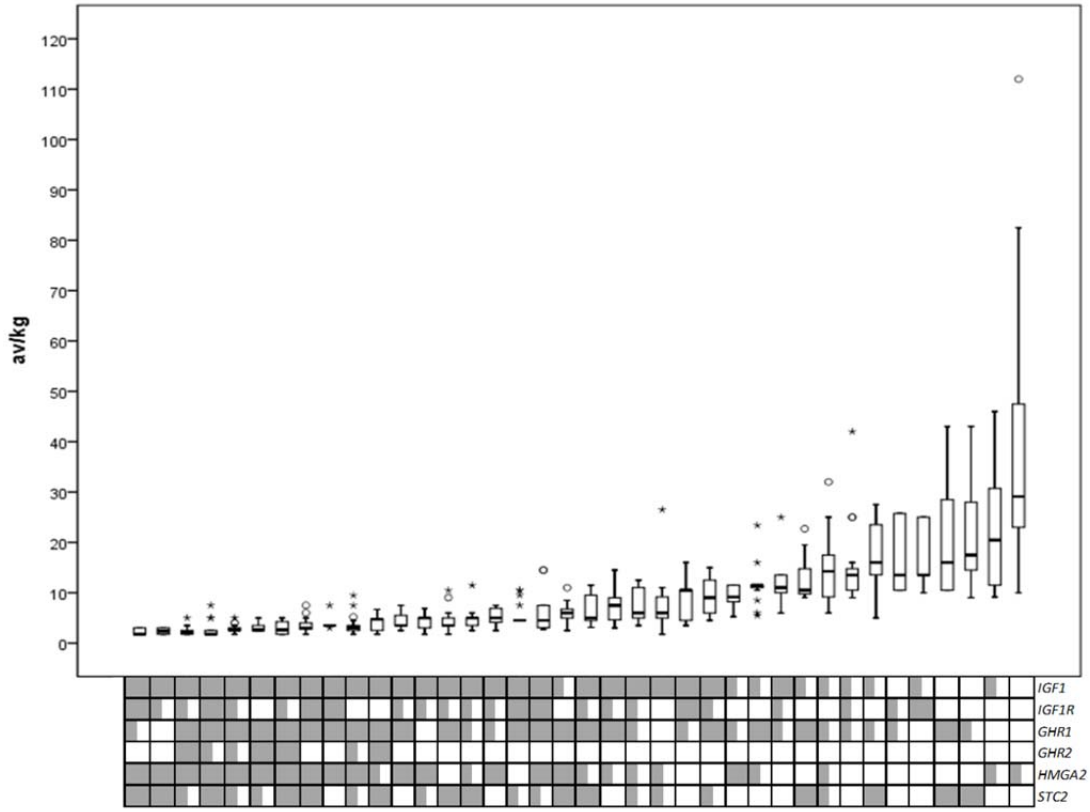

Figure 9: The weight of the dogs with the combinations found in 12–28 carriers. Derived markers marked in grey, ancestral markers in white.

As seen in the figure 9, the rare combinations result in smaller sized dogs than the most common combinations. The results of weight give a more liner outcome because the breed-defining chondrodysplasia may reduce the height of the shorter breeds.

The rough reducing effect of each allele is best seen in a graph of only homozygous alleles, in where all heterozygous alleles are considered derived homozygous alleles. The figures of the homozygous combinations in figure 10 and figure 11 show the reducing effect of each of the 41 homozygous allele combinations. The more derived alleles (marked in grey), the smaller the dog size is in both weight and height.

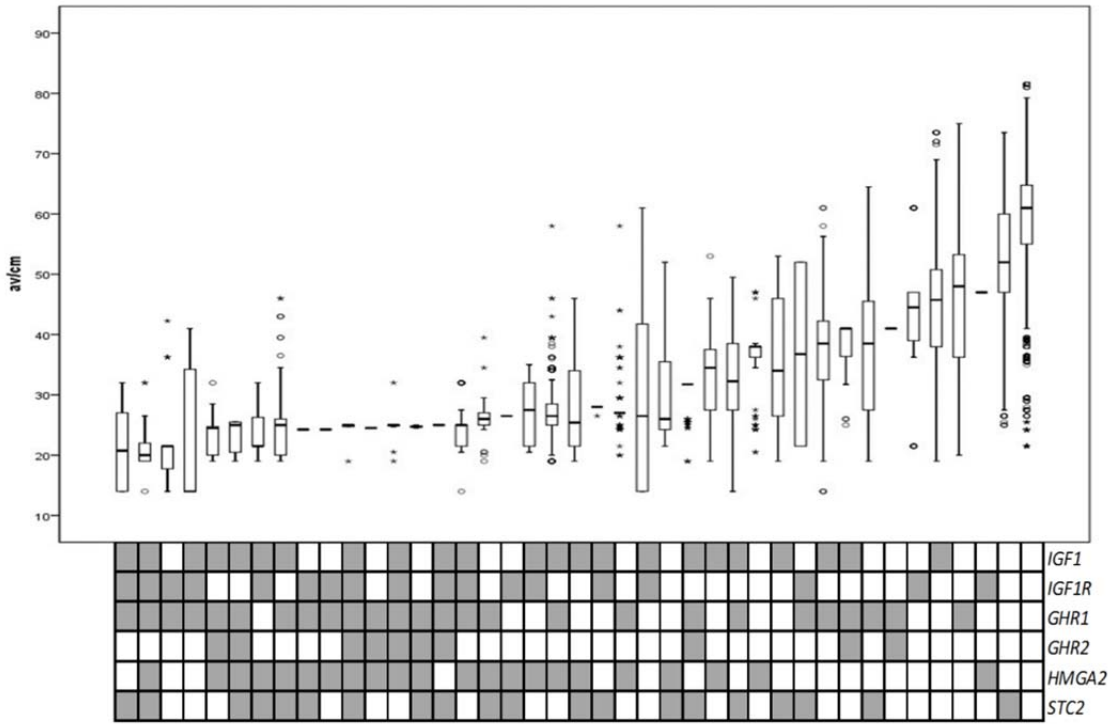

Figure 10: Height distribution of homozygous combinations. Derived markers marked in grey, ancestral markers in white.

The combinations are arranged according to the median of the size range of the combination. A close examination of the figure 10 may reveal that the higher number of derived alleles may not predict a smaller size. However, the breed defining chondrodysplasia may affect the results of especially the smaller sizes. In addition, the graph hides the effect of the heterozygous alleles, making the size range of each combinations wide. It also must be noted that the number of the breeds and individuals is not balanced. Some results may be due to small number of individuals especially within the breeds that have a small size range in the box plots.

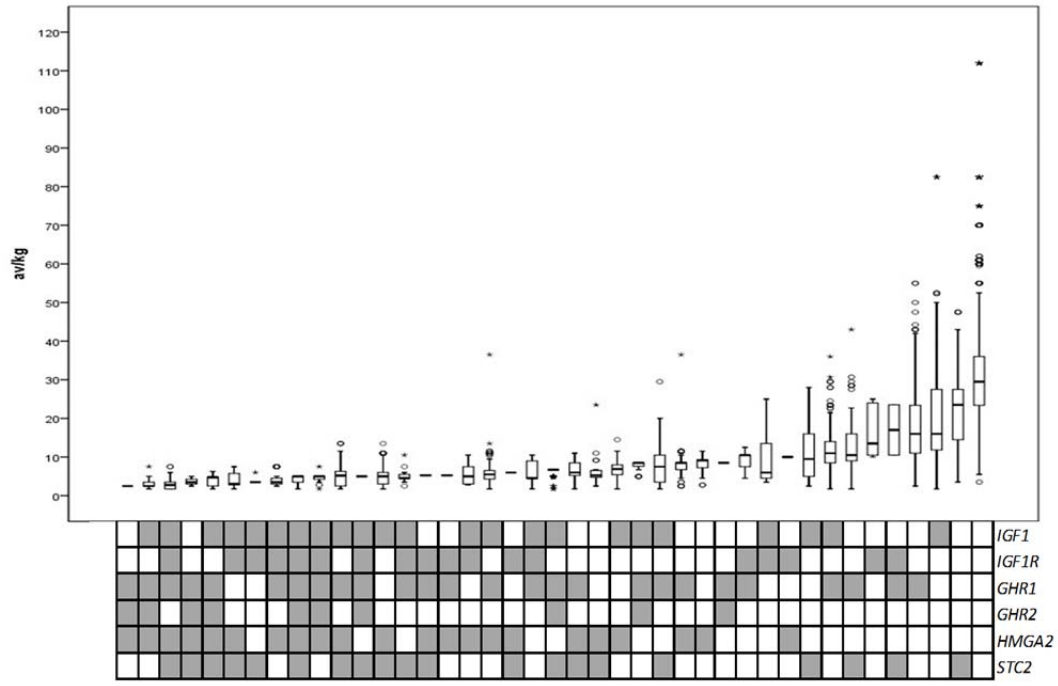

Figure 11: Weight distribution of homozygous combinations. Derived markers marked in grey, ancestral markers in white.

As seen in figure 11, which shows the weight distribution of the homozygous combinations, the combinations with none or only one derived allele results in the largest body size. Within the largest breeds, it may be found that the derived alleles of *IGF1*, *GHR1* and *STC2* are commonly found in the combinations. If the derived alleles of *HMGA2, GHR2* or I*GF1R* are in the combination, the size is remarkably reduced.

The comparison of any combination graph it is seen that the same combination does not contribute the same size in height or weight, possibly due to the short-legged breeds. The order of the combinations may vary when comparing the height and weight graphs of the homozygous combinations.

The combinations were suggested to be the most informative source for size prediction by Rimbault et *al.* (2013), and again as seen in the graphs of this study, the combinations indeed show the most detailed size variation within breeds.

6.5 Breed-defining chondrodysplasia

The aim in studies of the breed-defining chondrodysplasia (BDC) was to define the genetic markers behind the BDC, verify the breeds that are fixed to the BDC and possibly discover new chondrodysplastic breeds. In addition, the effect of the six canine size associated markers within the chondrodysplastic breeds were evaluated.

In total 16 372 dogs were tested for the three required genetic markers to qualify for the breed defining chondrodysplasia (BDC). There were 16 breeds found fixed to the breed-defining chondrodysplasia. Taken into consideration the potential genotyping errors, and other possible yet unknown genetic markers, breeds that had a frequency of BDC above 0.97 were also qualified for the breed defining chondrodysplasia status.

However, breeds with a BDC frequency below 0.97 were disqualified from the full status of the breed-defining chondrodysplasia but still kept in the analyses as it could be seen that some of these breeds may have shorter legs relative to their body size and in comparison with the controls.

The logic behind the test for the BDC status is introduced in the materials and methods section. In total, 1970 individuals from 40 breeds were found with some frequencies of chondrodysplasia. As many controls (1970 individuals from 40 breeds) were randomly chosen from the non-BDC breeds.

In Figure 12, the three-way test identifies the breeds with the BDC status by comparing the cases and controls. The previously identified alleles that result in shorter legs, thus shorter body, are the markers for the BDC. The graph identifies the rules for the three-way test: the BDC breeds likely have the *fgf4* retrogene insertion, and in addition also the A/G genotype in SNP1 and T/T in SNP2. As mentioned before, all of the three criteria must be fulfilled in order to qualify for the BDC. As seen in the graph of figure 12, A/G in SNP1 or T/T in SNP2 results in normal sized individuals if all the three rules are not filled.

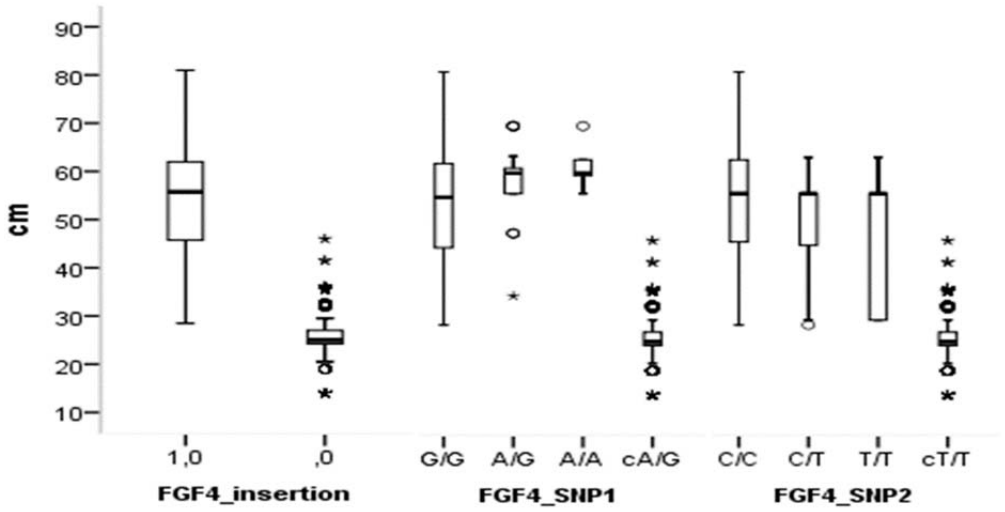

Figure 12: The breed-defining chondrodysplasia markers in comparison to the normal legged markers. Letter “c” indicates the chondrodysplasia status in breeds that have all the markers of the BDC.

Some breeds have been verified as BDC breeds in previous research by having a BDC frequency of 1.00 (Parker et *al.* 2009). The breeds that were verified by Parker et *al.* (2009) and again the results repeated in this study were: Cairn Terrier, Dachshund, Dandie Dinmont Terrier, Scottish Terrier, Skye Terrier, Swedish Vallhund, Lancashire Heeler, Norwich Terrier, Pekingese, Petit Basset Griffon Vendeen, Welsh Corgi Cardigan, Welsh Corgi Pembroke, West Highland white terrier and Basset hound.

In figure 13, the height of the normal legged controls is compared with the chondrodysplastic cases. The height range within the BDC dogs is not as wide as within the normal legged controls. However, similar, yet not equal trends are found to be caused by the derived markers in both the controls and cases. The influence of the derived alleles cause smaller individuals within all markers but surprisingly with the BDC breeds the derived *GHR1* allele results in larger individuals than the heterozygous allele. At least within the *GHR1* marker, most alleles are fixed for either derived, ancestral or homozygous allele within each breed. However, the distribution of the derived *GHR1* allele is not fixed to small breeds within the BDC status nor the ancestral allele to large breeds.

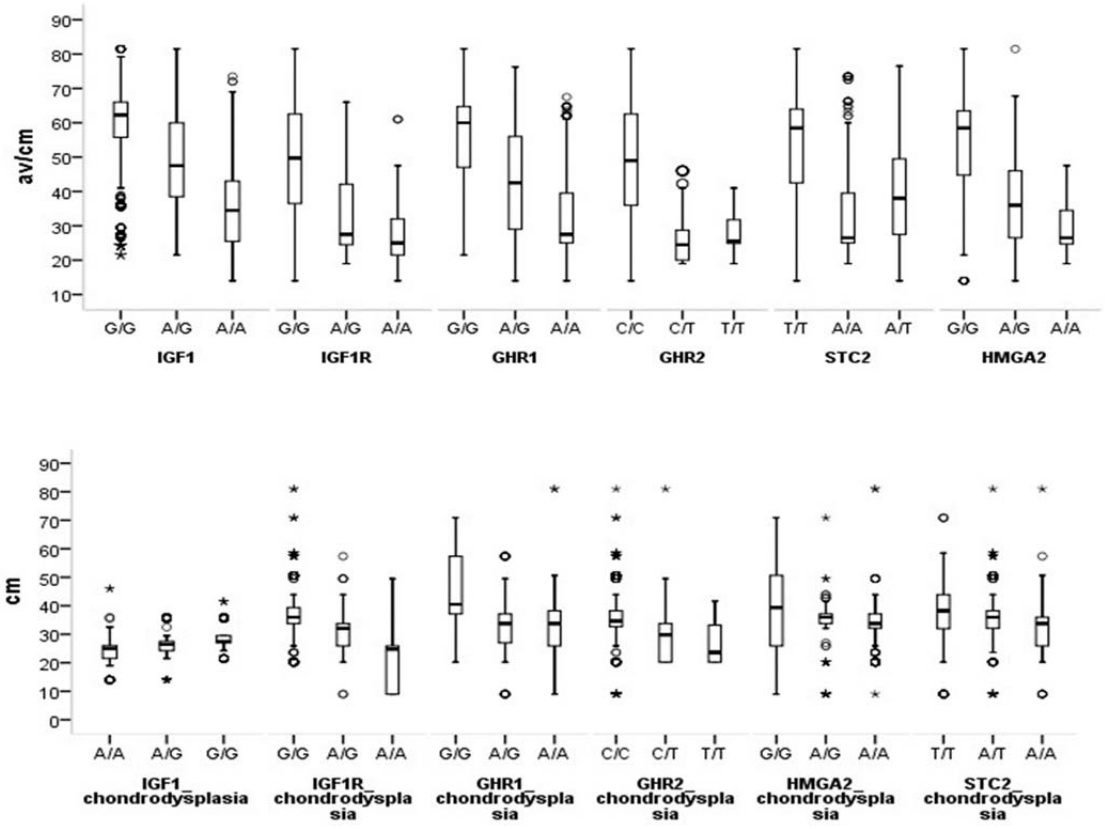

Figure 13: Comparing the six markers of the BDC breeds and normal legged breeds

Newly discover breeds in this study, that have not been previously verified genetically for the BDC are Australian Terrier (frequency 1.00), Norfolk Terrier (frequency 1.00), Sealyham Terrier (frequency 1.00), Cesky Terrier (1.00) Miniature Dachshund (frequency 0.99), Australian Silky Terrier (frequency 0.97), and Maltese (frequency 0.97). The phenotype of all of these breeds is indeed short-legged.

In the seven Jack Russell Terriers tested in the previous study by Parker et *al.* (2009), none had BDC. Surprisingly, in the data for this current research, out of the 84 Jack Russell Terriers tested, up to 45 individuals qualified for the BDC, giving the frequency of 0.54. The different results may be due to the different size data. Looking at the breed history of the Jack Russell Terrier, which was bred only for the working purposes to hunt fox, the small size became convenient in following the fox to the small holes. Bred most likely from the Fox Terri-

ers, Jack Russell Terriers have shorter legs and a slightly longer body than height, which allows them to be excellent working dogs. As the breed is developed for working, the breed standards regarding size are wide, which may explain that some strands of the Jack Russell terrier carry the BDC and so have shorted legs than the ones from the same breed with no BDC. (Jack Russell Terrier Club of America, Federation Cynologique Internationale.)

Irish Glen of Imaal Terrier is an old breed that had a breed standard published by the FCI only in 2001 and lately has been recognized by the American Kennel club in 2004 and (American Kennel Club, Federation Cynologique Internationale). The appearance of the breed is short legs and a large body, a phenotype of the BDC. (Federation Cynologique Internationale). However, in the data of 13 Irish Glen of Imaal Terriers tested, the frequency for the BDC was only 0.46. In the previous research by Parker et *al.* (2009) the frequency of the 12 tested Irish Glen of Imaal Terriers was 0.92. Parker et *al.* (2009) recalled that the low frequency of BDC for the obviously short legged breed was due to the recent recognition of the breed and hypothesized that the frequency would eventually reach 1.00. It is possible that the number of only 600 to 800 registered Irish Glen of Imaal Terriers in the US (Glen of Imaal Terrier Club of America) have not yet reached the complete genetic stability within the breed or the breed may have an additional, yet unknown genetic cause for the short legged phenotype.

Grand Basset Griffon Vendeen is the larger breed of the Petit Basset Griffon Vendeen, both coming from the same ancestor. Within the 11 tested Grand Basset Griffon Vendeens three were found to have the BDC. The frequency for the BDC in Petit Basset Griffon Vendeen is 1.00. Grand Basset Griffon Vendeen is a new breed in the US and has been recognized by the AKC only from 2014 (American kennel club), it is likely that the BDC in Grand Basset Griffon Vendeen will disappear in the next generations.

The Bichon group, originated from the Mediterranean small white dogs “Melitensis” includes Bolognese, Maltese, Havanese, Bichon Frise, Coton de Tulear and Lowchen (American Bolognese Club). The frequency for BDC in

Coton de Tulear is 0.89, for Bolognese 0.80, Bichon Frise 0.74, Maltese 0.97 and Havanese 0.91, Lowchen was not found with BDC.

The breed standard for Maltese and Bolognese has been valid only since the beginning of 2016 allowing tolerance for height from 20 cm to 25 cm. Havanese has been recognized by the FCI since 2009, with the tolerance for height 21 cm – 29 cm. Breed standard of the Bichon frises, published by the FCI in 1998, allows all height variation up to max 30 cm and the breed standard for the Coton de Tulear, published in 2000, allows variation in height from 22 cm to 30 cm. (Federation Cynologique Internationale). While the breed standards for the Bichon group are fairly new, the allowance for height variation for a small sized dogs is wide. Possibly, as long as there may be up to one third of the height variation within a breed, there is a chance that the frequencies for the BDC may never be fixed in these breeds.

Breeds with an origin in Tibet, such as Tibetan spaniel, Shih Tzu and Lhasa Apso (Federation Cynologique Internationale) were found with similar frequencies for the BDC, 0.95, 0.95 and 0.94 respectively. As these breeds are old, the frequencies for the BDC are expected to be stabilized, and the BDC will probably not be fixed within these breeds. In the previous research for the BCD by Parker et *al.* (2009) Tibetan Spaniel and Shih Tzu were found to be fixed for the BDC. However, the difference may be explained by the data size, as Parker tested only 10 Shih Tzus and 4 Tibetan spaniels, while in this research 57 Tibetan spaniels and 74 Shih Tzus were tested.

The Miniature Portugese Podengo is one of the three size variations of the Portugese Podengos. The frequency for the BDC in the 35 tested Miniature Portuguese Podengos was 0.69. Even though the breed is old, it has gotten popular only in past decade possibly the smallest variate has not had the time yet for the completely fixed BDC.

For the normal legged small breeds with some frequencies for BDC Parker et *al*. (2009) suggested that some additional growth factors may hide the BDC

from the phenotype and so allowing it to segregate within the breed. This might be the case for Yorkshire terrier (frequency 0.56) and Mi-Ki (frequency 0.7).

Chihuahua got the frequency rate of 0.7 out of the 20 individuals tested. The results show a higher breed-defining chondrodysplasia rate for the Chihuahua, than the research of Parker et *al.* 2009, yet still the possible additional markers for the small sized breeds may effect it.

Cesky Terrier has been created by crossbreeding the Scottish terrier and the Sealyham terrier in the 1950's (The Club of the Cesky Terrier Breeders). Both Scottish terrier and the Sealyham terrier are fixed for BDC, surprisingly however the Cesky terrier has a frequency of 0.7 for the BDC. However, the number of tested individuals was low. Ten tested individuals may not give a scientific proof of the status of the BDC in the breed and further research is required.

Biewer Terrier is a relatively new breed, only two decades old, (American Kennel Club) and has not been recognized by the FCI or the AKC. The fact that it has no breed standards, will explain the BDC frequency of 0.5.

For the low frequencies of BDC in the miniature Poodle (0.13) Rat Terrier (0.07) and the English Springer Spaniel (0.01) the possible cause is in the genotyping. These breeds show no signs of dwarfism in their phenotype, but further research may be required.

All breeds with any frequencies of the breed defining chondrodysplasia were taken in consideration as the cases for building the chart in figure 13 to compare the height of BDC with the height of the controls with similar weight individuals.

By comparing the controls (green dots) and the cases (red dots) in figure 14, the breed defining chondrodysplasia caused by *fgf4* retrogene was found to reduce the height of the individual by approximately 1.4 times to the normal control of the same weight. For breeds with similar size to the Norwegian Elkhound, this would make about 10 cm difference. In previous researches, similar results

about Norwegian Elkhounds height reduction caused by chondrodysplasia were accomplished. (Kyöstilä et *al.* 2013.)

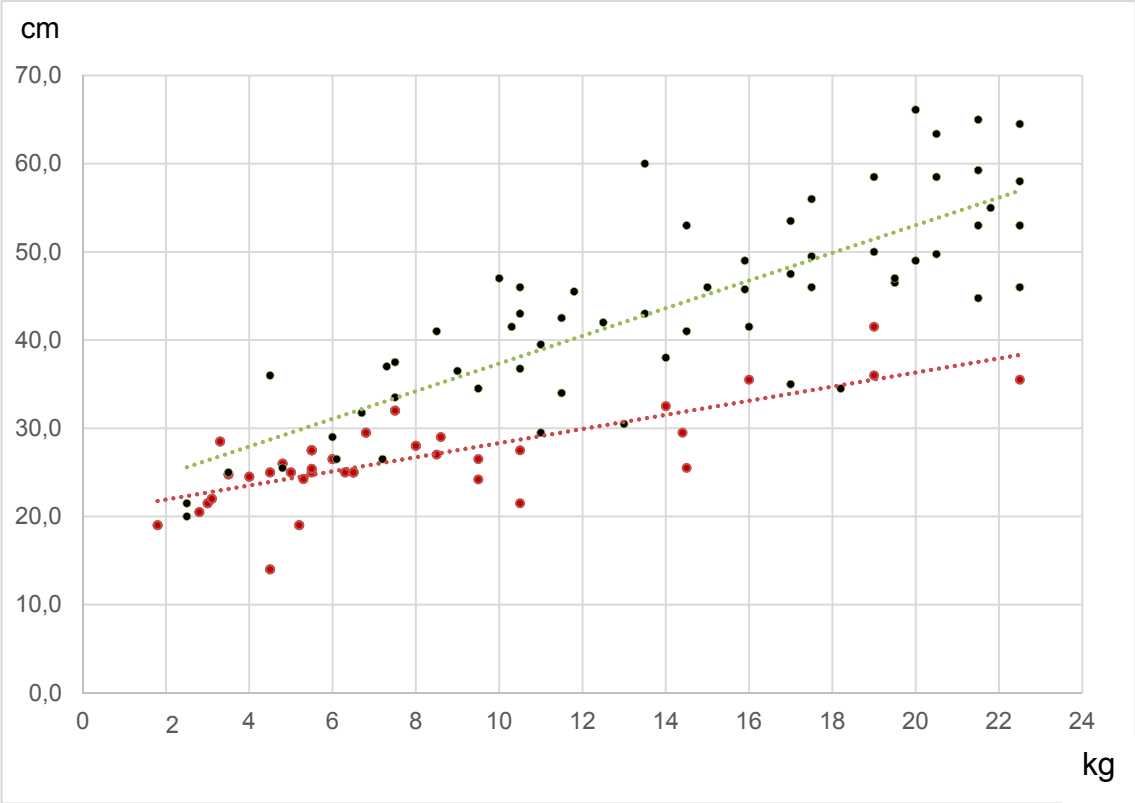

Figure 14: The height of breeds with breed-defining chondrodysplasia (red dots) compared with the height of normal legged breeds (green dots) in the same weight groups

Breed defining chondrodysplasia is one of the most straight forward genetic markers, giving a simple positive or negative result, which will either reduce height or not. In this study, about 20 breeds were verified to have BDC, of which seven breeds were newly-discovered. The breeds with chondrodysplasia can be identified genetically with three criteria, and the normal legged breed are about 1.4 times larger in height than the breed with BDC.

7 DISCUSSION

In this study the results of previous studies were repeated and new discoveries were made. In the statistical analyses it was remarked, that the derived alleles of the six markers previously identified by Rimbault et *al* (2013) reduce the size of dog and it was verified that the more there are derived alleles, the more the size is reduced.

Similarly to the results of Rimbault et *al.* (2013) it was found that canine with homozygous derived allele in any of the six loci had mean SBW from four to eight kilo, except for the derived *IGF1* allele*,* which had a wide mean weight distribution. The height of these breeds with homozygous derived allele in any of the six loci was below 30 cm, again, except for the derived *IGF1* allelle which could be found in most height ranges*.* As the derived *IGF1* allele is present in all weight and height categories, it indicates that the derived *IGF1* allele may not reduce the body size of canine if it is the only derived allele out of the six markers.

Even though the derived *IGF1* allele frequency is visible in all height and weight groups, the derived *IGF1* allele is most commonly found within the smaller breeds. The *IGF1* marker is straight forward with the height, however only if the reducing effect of the chondrodysplasia is taken into account. The frequencies of the derived alleles of *IGF1* in the weight categories is more complex because the derived *IGF1* allele seems to be found in most of the weight groups, almost as commonly in the heavier dogs as in the smaller breeds. Similar finding were made by Rimbault et *al.* (2013) who hypothesized that possibly the heavier breeds have additional markers that reduce the effect of the derived *IGF1* allele.

Derived *IGF1* allele however has some size reducing effect and in fact, it may explain the smallest size of the Entlebucher Mountain Dog (24.5 kg, 46 cm) within the Swiss mountain dogs. It may be hypothesized that the reducing effect

of fixed derived *IGF1* allele on a breed is strong, yet still its effect on individual size needs further studies.

Even though the derived *IGF1* allele has been strongly associated with reduced body weight in previous articles (Rimbault et *al.* 2013, Sutter et *al.* 2007), all weight categories from small to large have some presence of the derived *IGF1* allele or the derived *GHR1* allele. Because these alleles are present in all size categories, it indicates that they alone, with no other derived markers, are not able to explain size variation within breeds. In fact, when examining the combinations, this statement seems to be found true, indeed the combination of all ancestral alleles with a derived allele in *IGF1* results in size range from small to large, including 5 dogs weighing above 80 kg. In comparison, if *GHR1 i*s the only derived allele with all other alleles ancestral, the canine would weigh from 10 to 55 kg. It was also found that the derived *HMGA2* allele accompanied by derived *IGF1* allele have the most significant effect in reducing size compared to any other marker pair found in the most frequent combinations.

It has been suggested that the role of derived *IGF1* allele is more important in reducing the body size than *IGF1R*. In fact, it has hypothesized that *IGF1R* mutation has a modest effect on the size of canine. (Hoopes et al. 2010.) However, the deal with *IGF1R* seems to be more complex. According to the six markers boxplots in previous sections, (figure 1 and figure 2), the size of canine was reduced with derived *IGF1R* allele. In the combinations graphs, when examining only derived *IGF1R* allele with derived *IGF1* allele or derived *GHR1* allele, the reduction of size is smaller than predicted by the six markers. The heterozygous *IGF1R* alone in an ancestral combination may not reduce the mean weight, in fact, the weight distribution of that combination is from 11.5 kg in the Shetland Sheepdog to 40 kg in the Akita. Possibly both derived *IGF1* and derived *IGF1R* alleles do not reduce size significantly when not accompanied by other derived alleles. However, the fact that derived *IGF1R* allele cannot be found among the most common combinations, possibly indicates that it is a relatively rare allele. As the least frequent combinations are also smallest in size, the impression may be that the rare, derived *IGF1R* allele is strong in reducing the canine size.

In the statistical analyses of the boxplots, it was noticed that the derived *IGF1R* allele and the derived *GHR2* allele were the only nonsignificant pairs when comparing the heterozygous derived allele with the homozygous derived allele. The results by Rimbault et *al.* (2013) give the same outcome for the markers in question. The reason for this may be that the impact of the derived *IGF1R* and the derived *GHR2* variants on the phenotype are so strong that even only a single derived allele reduces the size of canine greatly. (Rimbault et *al.* 2013.). This however does not explain the fact that if derived *IGF1R* is the only derived allele in an ancestral combination, the size of dog is not reduced. Possibly the strong reducing effect of derived *IGF1R* allele, even with just one copy, is only visible when other markers are present.

Rimbault et *al*. (2013) predicted *HMGA2* and *GHR2* to be the most dominant markers. However, in addition to *GHR2* and *HMGA2*, also *IGF1R* seem to have a very a strong derived allele, which size reducing effect may be distinctly seen in the phenotype. Even though these markers mostly may require the presence of additional size reducing marker in the combination, the frequencies of these alleles seem to be very low or infrequent in canines weighing above 8 kg. The only two ancestral alleles that are fixed in larger breeds alone are *IGF1R* and *GHR2*. Therefore these two markers seem to have the strongest impact on reducing the body size.

One interesting observation was that 43% of the breeds that carry one copy of the derived *GHR2* allele are short legged breeds. This adds up to the total of 33% of the total number of the dogs that are carrying the heterozygous allele in *GHR2,* are short in height also due to some other markers, such as *fgf4*. Half of the six breeds with a fixed derived *GHR1* allele are breeds with chondrodysplasia, which may explain the small height within the breeds with a fixed *GHR1* and chondrodysplasia. The reason for the derived *GHR1* commonly being accompanied by chondrodysplasia is yet unknown. Rimbault *et* al. (2013) stated that *GHR1* and *GHR2* are most likely inherited together, the chondrodysplasia with derived *GHR2* may be explained by the fact that derived *GHR2* is found only if *GHR1* is also derived. This however does not explain why the chondrodysplasia

is more commonly found within the dogs with derived *GHR.* In addition, only nine breeds were found with homozygous derived *GHR2* allele. Two of the breeds were short legged breeds with chondrodysplasia, while 12 out of the 26 breeds that had heterozygous *GHR2* allele, had also breed-defining chondrodysplasia. However, very surprisingly *GHR2* heterozygous allele produces canine with the mean weight and height lower than the homozygous derived allele. The mean weight of the homozygous *GHR2* is 5.2 kg and mean height 24.5 while the mean weight for heterozygous *GHR2* is 4 kg and mean height 25.5 cm. Neither heterozygous nor homozygous derived *GHR2* allele have been found fixed in any breed. Possibly the chondrodysplasia status may be one of the explaining matters for the fact that breeds with heterozygous *GHR2* allele results in smaller mean body size or at least shorter height.

One other promising marker in canine size reduction is the *STC2* marker. Because the heaviest breeds with a fixed derived *STC2* allele are still tiny dogs, it is most likely that the derived *STC2* allele has a strong size reducing effect when fixed to the breed. Homozygous derived *STC* allele was found to have a significant size reducing effect, but the effect was weak when only one copy of derived *STC* allele was present. However, with the presence of derived *IGF1* allele the size was considerably reduced. The fact that the heterozygous *IGF1* allele and heterozygous *STC2* allele reduce the size only weakly, if at all, indicates that the heterozygous derived pair of derived *STC2* allele and derived *IGF1* allele amplify each other effects, making them strong together. Another interesting notion is the unpretentious nature of derived *STC2* allele; if there is more than 3 derived markers in the combination, the size reducing effect of derived *STC2* gets hidden.

As seen in figure 6 and figure 7 in the previous section, within all the combinations the size is reduced similarly in both weight and height. It seems, that in the most common combinations, excluding the *GHR2*, the most important factor for the canine size is the number of the derived alleles, especially when there are only two derived alleles in a combination. Only secondly the identity of the alleles in question matter, mostly in the determining of size in the fine scaling.

A rough guide line for predicting the frequencies of the combinations may be that the more there is derived alleles in the combination, the less there is carriers of that combination. In addition, it was discovered that the rarest combinations were also carried by the smallest breeds. The distribution of the combinations within the smaller breeds seems to be much wider than within the larger sized dogs might be because the larger breeds carry more ancestral alleles, while the smaller breeds also carry a variety of derived alleles, allowing more variation in the combinations.

Only a few breeds were found fixed to a combination, but no combination was fixed to a breed. Similarly to the result by Rimbault et *al.* (2013) it was observed that most combinations were not fixed to any breed and there were typically multiple combinations present within a breed. In fact, only two combinations were found fixed to a breed. Yet still, the combinations of alleles are believed to explain size variation in dogs better than the individual markers. (Rimbault et *al.* 2013.)

The most common combination was the combination with all ancestral alleles was found in breeds that are considered ancestral-like. The combination was also one of the two combinations fixed to some breeds. These breeds with all ancestral combination were found to be wolf-like in size, with SBW of 52-60 kg.

In comparison, no breed was found fixed to the combination with all derived alleles and in fact the combination was found to be carried by only 10 dogs, all of them considered tiny breeds. However, the derived alleles were not rare at all and they were oddly found in all size categories. The fact that even the largest breeds carry a fair amount of derived alleles may verify the hypothesis by Rimbault et *al.* (2013) that the larger breeds must have additional markers making them larger than their ancestors. The size group with the least derived alleles was the third largest group, considered wolf-sized. On the other hand, the group that is the smallest in size that carries the highest number of fixed ancestral alleles (which include *IGF1R, GHR2,* and *HMGA2*) is weighing 22 kg.

The fixed alleles were found to be an informative source for examining size differences between the breeds, but it is not reliable enough for the use in individual scale as the dog size within the breed may vary.

The most important findings were the relationships of the six genetic markers as well as the identifying of the breeds with chondrodysplasia. The frequency tables and the examining of the fixed alleles and the box plots gives a valuable information on the overall size range of the breed. Examining the combinations is the most efficient method to determinate the individual size difference as well as the effect of each marker. An interesting aspect for the future would be the studies on whether the six markers predict smaller individual body size within the breed.

8 CONCLUSIONS

The aim of this study was to examine the possibility of predicting the genetic body size of dog. The ability to predict the genetic size of a dog would ease the creation of a size associated diet and exercise for each individual, which would possibly result in a healthier and longer life. The size associated qualities may facilitate the decision of the right individual for each size associated purpose Also size prediction may allow the choosing and training of the right sized individuals, already at puppy age, for different hobbies such as agility, dog races or dog shows, which may have very strict size regulations

In this study the findings of Rimbault et *al.* (2013) on the dog size were replicated. The previously found markers were examined according to the association of individual markers with size, the frequencies and the alleles fixed to a breed as well as the association of the multimarker combination with size. As expected, the size associated markers were found to reduce the dog body size. In addition, the findings by Parker et *al*. (2009) were replicated and some previously found chondrodysplastic breeds were verified and new short legged breeds discovered. The chondrodysplasia status was evaluated with the size associated markers and finally, the customer report view was developed according to the findings.

The studies on the size associated markers reveal that, the more there is derived alleles in the combination, the more the size of the dog is reduced. In addition to the size reducing effect of the homozygous derived alleles, some dominant alleles reduce the dog size significantly even with only one copy, especially when accompanied by other derived markers. On the other hand, the effect of some heterozygous markers may not even be seen in the phenotype, even if accompanied by other derived alleles.

In general, the more there is derived alleles in the size associated genetic marker combination, the smaller is the body size of a dog and most importantly the number of the allele's matters instead of the identity of the allele. However,

some alleles seem to be stronger in size reducing than other alleles. The derived alleles that are considered to have a stronger size reducing effect are *GHR2*, *HMGA2* and *IGF1R*. The alleles that in this study are considered to have a weaker size reducing effect are derived *IGF1*, *GHR1* and *STC2.*

Mostly, the alleles do not reduce the size of the dog significantly if they are the only derived alleles in the ancestral combination. Out of the individual alleles in combination, the derived *HMGA2* allele seems to have the strongest size reducing effect when it is the only derived allele in a combination. The ancestral combination predicts the size of the dog in the size range of a wolf, but if there is derived homozygous *HMGA2* allele present in ancestral combination, the median weight is about 10 kg. When the only derived allele is heterozygous *HMGA2,* the size of canine is not remarkably reduced

Out of the individual allele's one of the strongest size reducing heterozygous alleles seem to be in derived *GHR2* but possibly because it is only found accompanied by at least one other derived marker. The derived GHR2 allele is in fact only found accompanied by the derived *GHR1* allele and never alone in an ancestral combination. Yet still, because the derived *GHR2* allele is present only in tiny dogs, it seems to have the strongest size reducing impact out of all markers when present in a combination compared to any other marker. In addition, the heterozygous *GHR2* marker was found in canine with the mean weight and height lower than the homozygous derived *GHR2* allele. In addition, the heterozygous *GHR2* allele seems to be accompanied by the chondrodysplasia status in nearly half of breeds with *GHR2*, which is over one third of the dogs with one copy of the derived *GHR2* allele. The reason for this is yet unknown but is possibly due to the crossbreeding of small and short-legged dogs in the early times of the modern dog breed formation.

As derived *IGF1* and *GHR1* alleles are found in all height and weight categories with high frequencies, among the larger breeds as well as within the small breeds, it may be possible that they do not reduce the size of the dog unless accompanied by some other derived alleles. Another explanation may be what has been proposed previously by Rimbault et *al.* (2013) that the larger dogs

may have other markers that hide the effect of derived *IGF1* allele and possibly also the effect of derived *GHR1* allele. The derived *GHR2* allele was found only within tiny breeds and small size categories. Derived *IGF1R* was found alone as the only derived allele in a few larger breeds and also heterozygous *HMGA2* is can be found in some larger sized canine.

The rarer the derived alleles is the stronger is size reducing effect of the allele. The rarest allele found was the derived *GHR2* allele, which was also found to have the strongest size reducing effect of all alleles, even if the derived *GHR2* allele cannot be found as a single derived allele in an ancestral combination. Another rare and strong size reducing alleles were found to be the derived *HMGA2* allele and possibly also the derived *IGF1R* allele. Within these markers only a few breeds were found with a fixed derived allele while mostly the breeds were fixed to the ancestral allele. The most common derived alleles are found in *IGF1, GHR1* and *STC2,* and because these alleles were found in most size categories, they seem to have more moderate effect on size than the mentioned rare alleles, Even though the derived *IGF1* allele or derived *GHR1* allele are considered to reduce the size of dog mildly, the effect of these allele in the individual size needs further studies and in fact they may explain the size variation within the breeds.

As suggested previously by Rimbault et *al.* (2013), the combinations are the most informative source for size prediction of canine. The most reliable size prediction can be given based on rare combinations with multiple derived alleles. The more there is derived alleles in a combination, the more the size of the dog is reduced. The most important factor in the size prediction seems to be the number of the derived alleles instead of the identity of the derived allele.

Most of the combinations are found across the breeds and only a few breeds are fixed to all same alleles. Similarly, no height or weight category has any fixed derived alleles but some ancestral alleles are fixed to most of the larger size categories. Knowing the nearly fixed alleles in the breeds, may help when determining the individual body sizes specially when examining the exceptions to the nearly fixed alleles of the breed.

With the results displayed in this research, a rough dog size prediction is possible by creating expected size ranges for each allele combination. For some of the combinations, mostly within the smaller breeds, the size prediction may be specific with the markers examined in this research. However, with the knowledge available at the moment, within the larger breeds the only reliable statement concerning size is that the dog is not a small or medium sized dog but the true weight and height of a large dog cannot yet be predicted.

However, during the preparation of this study an additional eleven new markers were discovered and published. These markers are hypothesized to predict nearly 90 % of the size of the dog. (Heyward et *al.* 2016). Future studies will possibly include these additional markers along with the markers already studied. Furthermore, despite the fact that the number of individuals and breeds was the largest ever studied for the size of a dog, a larger number of breeds and dogs would predict a more reliable outcome.

Eventually, the results of this study could be applied into a mathematical theorem which would allow the prediction of dog size. At the moment, the approximate prediction for the customer view report is achievable, which allows the possibility to collect individual sizes for further analyses.

REFERENCES

Alberts, B.; Johanson, A.; Lewis, J.; Raff, M.; Roberts, K.; Walter, P. 2008. Molecular Biology of the cell. Reference Edition. 5th Edition. Garland Science. United States of America.

American Bolognese Club. http://americanbologneseclub.com/breed-info/history/. Referred 14.02.2016.

American Kennel Club. http://www.akc.org/. Referred: 10.2.2016.

American MI-Ki registry association. http://www.americanmi-kiregistryassociation.com/index.php?id=standard. Referred 21.1.2016.

American Rare Breed association. http://www.arba.org/. Referred: 17.1.2016.

Australian Bull Arab registry Inc. http://www.bullarab.com.au/bull-arab-standard.html. Referred 17.2.2016.

Australian labradoodle association Inc. http://www.laa.org.au/index.php/about-labs/standards. Referred 17.2.2016.

Beer A. 2005. Maailman nisäkkäitä. Suom. Kalliola, I. Animal fact file. Otava, Helsinki.

Biewer breed club of America. http://www.biewerbca.org/biewer-standard.html. Referred: 14.2.1.2016.

Boerboel International. http://www.boerboelinternational.com/en_US/. Referred 20.1.2016.

Boyko, A.; Quignon, P.; Li, L.; Schoenebeck, J.; Degenhardt, J.; Lohmueller, K.; Zhao, K.; Brisbin, A.; Parker, H.; vonHoldt, B.; Cargill, M.; Auton, A.; Reynolds, A.; Elkahloun, A.; Castelhano, M.; Mosher, D.; Sutter, N.; Johnson, G.; Novembre, J.; Hubisz, M.; Siepel, A.; Wayne, R.; Bustamante, C.; Ostrander, E. 2010. A simple Genetic architectute Underlies Morphological Variation in dogs. Plos Biology 8.

The Club of the Cesky Terriers. http://www.ceskyterier.cz/history.htm. Referred 14.02.2016.

Coursier, EA.; Thomson, RM.; Mellow, DJ.; Yam, PS. 2010. An epidemiological study of environmental factors associated with canine obesity. J Small Anim Pract. 7:362-7.

Dennis-Bryan, K. 2012. The Complete Dog Breed Book. Dorling Kindersley. London.

Din, Z.; Oskarsson, M.; Ardalan, A.; Angleby, H.; Dahlgen, L.; Tepeli, C.; Kirkness, E.; Savolainen, P.; Zhang, Y. 2012. Origins of domestic dog in Southern East Asia is supported by analysis of Y-chromosome DNA. Heredity 108, 507–514.

English toy terrier club. http://www.english-toy-terrier-club.co.uk/kennel-club-standard/. Referred 02.02.2016.

Federation Cynologique Internationale. http://www.fci.be/en/. Referred 18.02.2016.

Fodge, B. Dog: The Definitive Guide for Dog Owners. 2010. Octopus Publishing. London.

French Bulldog club of America. http://frenchbulldogclub.org. Referred 18.02.2016.

German, A. The growing Problem of Obesity in Dogs and Cats. 2006. American society for nutrition. vol. 136. no. 7: 19405-19465.

Glen of Imaal Terrier Club of America. http://www.glens.org/physical_description.html. Referred 14.02.2016.

Glickman, LT.; Glickman, NW; Schellenberg, DB.; Simpson, K.; Lantz, GC. 1997. Multiple risk factors for the gastric dilatation-volvulus syndrome in dogs: a practitioner/owner case-control study. Journal of the American Animal Hospital Association: 33(3):197-204.

Gray, M.; Sutter, N.; Ostrander, E.; Wayne, R. 2010. The IGF1 small dog haplotype is derived from Middle Eastern grey wolves. BMC Biology 8:16.

Hayward, J.; Castelhano, M.; Oliveira, K.; Corey, E.; Balkman, C.; Baxter, T.; Casal, M.; Center, S.; Fang, M.; Garrison, S.; Kalla, S.; Korniliev, P.; Kotlikoff, M.; Moise, N.; Shannon, L.; Simpson, L.; Sutter, N.; Todhunter, R.; Boyko, A. 2016. Complex disease and phenotype mapping in the domestic dog. Nature Communications, 7:10460.

Hock, R.; Furusawa, T.; Ueda, T.; Bustin, M. 2008. HMG chromosomal proteins in development and disease. Trends Cell Biol.; 17(2): 72–79.

Hoopes, B.; Rimbault, M.; Liebes, D.; Ostrander, E & Sutter, N. 2012. The insulin-like growth factor 1 receptor (IGF1R) contributes to reduced size in dogs. Genome Research 23: 1985-1995.

Institute of canine biology. http://www.instituteofcaninebiology.org/norwegian-lundehund.html. Referred 2.3.2015.

Irion, D.; Schaffer, A.; Famula, T.; Eggleston, M.; Hughes, S.; & Pedersen N. 2003. Analysis of Genetic Variation in 28 Dog Breed Populations with 100 Microsatellite Markers. Jounal of Heredity 94: 81-87.

Jack Russell Terrier Club of America. https://www.therealjackrussell.com/breed/history.php. Referred 14.02.2016.

Jepsen, M.; Kløverpris, S.; Mikkelsen, J.; Pedersen, J.; Füchtbauer, E.; Laursen, L.; Oxvig, C. 2014. Stanniocalcin-2 Inhibits Mammalian Growth by Proteolytic Inhibition of the Insulin-like Growth Factor Axis. JBC: The journal of Biochemistry.

Jones, P.; Chase, K.; Martin, A.; Davern, P.; Ostrander, E.; Lark, K. 2008. Single-Nucleotide-Polymorphism-Based Association Mapping of Dog Stereotypes. Genetics 179: 1033-1044.

Kretahund club. http://www.kretahund.de/kritikoslagonikos/000000a2d412f4403/standardenglish/index.html. Referred.16.02.2016.

Kyöstilä, K.; Lappalainen, A.; Lohi, H. 2012 Canine Chondrodysplasia Caused by a Truncating Mutation in Collagen-Binding Integrin Alpha Subunit 10. PLoS ONE 8(9): e75621.

Lindblad-Toh, K.; Wade, C.; Mikkelsen, T.; Karlsson, E.; Jaffe, D.; Kamal, M; Clamp, M.; Chang, J.; Kulbokas, E.; Zody, M.; Mauceli, E.; Xie, X, Breen, M.; Wayne, R.; Ostrander, E.; Ponting, C.; Galibert, F.; Smith, D.; deJong, P.; Kirkness, E.; Alvarez, P.; Biagi, T.; Brockman, W.; Butler, J.; Chin, C.; Cook, A.; Cuff, J.; Daly, M.; DeCaprio, D.; Gnerre, S.; Grabherr, M.; Kellis, M.; Kleber, M.; Bardeleben, C.; Goodstadt, L.; Heger, A.; Hitte, C.; Kim, L.; Koepfli, K.; Parker, H.; Pollinger, J.; Searle, S.; Sutter, N.; Thomas, R.; Webber, C.; Baldwin, J.; Abebe, A.; Abouelleil, A.; Aftuck, L.; Ait-zahra, M.; Aldredge, T.; Allen, N.; An, P.; Anderson, S.; Antoine, C.; Arachchi, H.; Aslam, A.; Ayotte, L.; Bachantsang, P.; Barry, A.; Bayul, T.; Benamara, M.; Berlin, A.; Bessette, D.; Blitshteyn, B.; Bloom, T.; Blye, J.; Boguslavskiy, L.; Claude Bonnet, C.; Boukhgalter, B, Brown, A.; Cahill, P.;16, Nadia Calixte16, Jody Camarata16, Yama Cheshatsang, Y.; Chu, J.; Citroen, M, Collymore, A.; Cooke, P.; Dawoe, T.;Daza, R.; Decktor, K.; DeGray, S.; Dhargay, N.; Dooley, K.; Dooley, K.; Dorje, P.; Dorjee, K.; Dorris, L.; Duffey, N.; Dupes, A.; Egbiremolen, O.; Elong, R.; Falk, J.; Farina, A.; Faro, S.; Ferguson, D.; Ferreira, P.; Fisher, S.; FitzGerald, M.; Foley, K.; Foley, C.; Franke, A.; Friedrich, D.; Gage, D.; Garber, M.; Gearin, G.; Giannoukos, G.; Goode,

T.; Goyette, A.; Graham, J.; Grandbois, E.; Gyaltsen, K.; Hafez, N.; Hagopian, D.; Hagos, B.; Hall, J.; Healy, C.;Hegarty, R.; Honan, T.; Horn, A.; Houde, N.; Hughes, L.; Hunnicutt, L.; Husby, M.; Jester, B.; Jones, C.;Kamat, A.; Kanga, B.; Kells, C.; Khazanovich, D.; Chinh Kieu, A.; Kisner, P.; Kumar, M.; Lance, K.; Landers, T.; Lara, M.; Lee, W.; Leger, J.; Lennon, N.; Leuper, L.; LeVine, S.; Liu, J.; Liu, X.; Lokyitsang, Y.; Lokyitsang, T.; Lui, A.; Macdonald, J.; Major, J.; Marabella, R.;Maru, K.; Matthews, C.; McDonough, S.; Mehta, T.; Meldrim, J.; Melnikov, A.; Meneus, L.;Mihalev, A.; Mihova, T.; Miller, K.; Mittelman, R.; Mlenga, V.; Mulrain, L.; Munson, G.; Navidi, A.; Naylor, J.; Nguyen, T.; Nguyen, N.; Nguyen, C.; Nguyen, T.; Nicol, R.; Norbu, N.; Norbu, C.; Novod, N.; Nyima, T.; Olandt, P.; O'Neill, B.; O'Neill, K.; Osman, S.; Oyono, L.; Patti, C.; Perrin, D.; Phunkhang, P.; Pierre, F.; Priest, M.; Rachupka, A.; Raghuraman, S.; Rameau, R.; Ray, V.; Raymond, C.; Rege, F.; Rise, C.; Rogers, J.; Rogov, P.; Sahalie, J.; Settipalli, S.; Sharpe, T.; Shea, T.; Sheehan, M.; Sherpa, N.; Shi, J.; Shih, D.;Sloan, J.; Smith, C.; Sparrow, T.; Stalker, J.; Stange-Thomann, N.; Stavropoulos, S.; Stone, C.; Stone, S.; Sykes, S.; Tchuinga, P.; Tenzing, P.; Tesfaye, S.; Thoulutsang, D.; Thoulutsang, Y.; Topham, K.; Topping, I.; Tsamla, T.; Vassiliev, H.; Venkataraman, V.; Vo, A.; Wangchuk, T.; Wangdi, T.; Weiand, M.; Wilkinson, J.; Wilson, A.; Yadav, S.; Yang, S.; Yang, X.; Young, G.; Yu, Q.; Zainoun, J.; Zembek, L.; Zimmer, A.; Lander, E. 2005. Genome sequence, comparative analysis and haplotype structure of the domestic dog. Nature. 438, 803-819.

McGreevy, P.; Georgevsky, D.; Carrasco, J.; Valenzuela, M.; Duffy, D.; Serpell, J. 2013. Dog Behavior Co-Varies with Height, Bodyweight and Skull Shape. PLoS One: 8 (12): e80529.

National Genome Research institute. https://www.genome.gov/10000533. Referred 18.02.2016.

Ortega-Pacheco, A.; Rodríguez-Buenfil, JC.; Segura-Correa, JC.; Bolio-Gonzalez, ME.; Jiménez-Coello, Forsberg, L.; 2006. Pathological Conditions of the Reproductive Organs of Male Stray Dogs in the Tropics: Prevalence, Risk Factors, Morphological Findings and Testosterone Concentrations. Reproduction in Domestic Animals: 41, Issue 5, pages 429–437.

Oskarsson, M.; Klütsch, C.; Boonyaprakob, U.; Wilton, A.; Tanabe, Y.; Savolainen. 2011. Mitochondrial DNA data indicate an introduction through Mainland Southeast Asia for Australian dingoes and Polynesian domestic dogs. The Roal Society Publishing.

Parker, H. 2012. Genomic Analyses of Modern Dog Breeds. Mammalian Genome 23: 19-27.

Parker, H, Shearin, A.; Ostrander, E. 2012. Man's Best Friend Becomes Biology's Best in Show: Genome Analyses in the Domestic Dog. Annual Review if genetics 44: 309-336.

Parker, H.; VonHoldt, B.; Quignon P.; Margulies E.; Shao S.; Mosher, D.; Spady, T.; Elkahloun, A.; GArgill, M.; Jones, P.; Maslen, C.; Acland, G.; Sutter N.; Kuroki, K.; Bustamante, C.; Wayne R.; Ostrander, E. 2009. An Expressed Fgf4 Retrogene is Associated with Breed-Defining Chondrodysplasia in Domestic Dogs. Science 325: 995-998.

Postel-Vinay, MC.; Kelly, PA. 1996. Growth hormone receptor signaling. Baillieres Clin Endocrinol Metab, 10 (3): 323-36.

Rimbault, M.; Beale, H.; Schoenebeck, J.; Hoopes, B.; Allen, J.; Kilroy-Glynn, P.; Wayne, R.; Sutter, N. & Ostrander, E. 2013. Derived variants at six genes explain nearly half of size reduction in dog breeds. Genome research, 1985-1995.

Rimbault, M. & Ostrander, E. 2012. So many doggone traits: mapping genetics of multiple phenotypes in the domestic dog. Oxford University Press.

Smith, GK.; Paster, ER.; Powers, MY.; Lawler, DF.; Biery, DN.; Shofer, FS.; McKelvie, PJ.; Kealy, RD. 2006. Lifelong diet restriction and radiographic evidence of osteoarthritis of the hip joint in dogs. J Am Vet Med Assoc.: 229(5):690-3.

Sutter, N.; Bustamante, C.; Chase, K.; Gray, M.; Zhao, K.; Zhu, L.; Padhukasahasram, B.; Karlins, E.; Davis, S.; Jones, P.; Quignon, P.; Johnson, G.; Parker, H.; Fretwell, N.; Mosher, D.;

Lawler, D.; Satyaraj, E.; Nordborg, M.; Lark, G.; Wayne, R. & Ostrander, E. 2007. A single IGF1 Allele Is a Major Determinant of Small Size in Dogs. Science. 316: 112-115.

Sutter, N.; Eberle, M.; Parker, H.; Pullar, B.; Kirkness, E.; Kruglyak, L.; Ostrander, E. 2004. Extensive and breed-specific linkage disequilibrium in Canis familiaris. Genome Research 14: 2388-2396.

UC Davis genome center. http://dnatech.genomecenter.ucdavis.edu/infinium-assay/. Referred 10.02.2016.

Vaysse, A.; Ratnakumar, A.; Derrien, T.; Axelsson, E.; Peilberg, G.; Sigarddsson, S.; Fall, T.; Seppälä, E.; Hansen, M ;Lawley, C.; Karsson, E.; Bannach, D.; Vila, C.; Lohi, H.; Galibert, F.; Fredholm, M.; Häggström, J.; Hedhammar, Å.; André, C.; Lindblad-Toh, K.; Hitte, C.; Webster, M. 2011.; Identification of Genomic Regions Associated with Phenotypic Variation between Dog Breeds using Selection Mapping. PLoS Genet.: e1002316.

Vilà, C.; Savolainen, P.; Maldonado, J.; Amorim, I.; Rice, J, Honeycutt, R.; Crandall, K.; Lundeberg, J.; Wayne, K. 2013. Multiple and Ancient Origins of the Domestic Dog. Science 276: 1687-1689.

vonHoldt, B.; Pollinger, J.; Earl, D.; Knowles, J.; Boyko, A.; Parker, H.; Geffen, E.; Pilor, M.; Jedrzejewski, W.; Jedrzejewska, B.; Sidorovich, V.; Greco, C.; Randi, E.; Musiani, M.; Kays, R.; Bustamante, C.; Ostrander, E.; Novembre, J. & Wayne, R. 2011. A genome-wide perspective on the evolutionary history of enigmatic wolf-like canids. Genome Research 21: 1294-1305.

vonHoldt, B.; Pollinger, J.; Lohmueller, K.; Han. E.; Parker, H.; Quignon, P.; Degenhardt, J.; Boyko, A.; Earl, D.; Auton, A.; Reynolds, A.; Bryc, K.; Brisbin, A.; Knowles, J.; Mosher, D.; Spady, T.; Elkahloun, A.; Geffen, E.; Pilot, M.; Jedrzejewski, W.; Greco, C.; Randi, E.; Bannasch, D.; Wilton, A.; Shearman, J.; Musiani, M.; Cargill, M.; Jones, P.; Qian, Z.; Huang, W.; Ding, Z.; Zhang, Y.; Bustamante, C.; Ostrander, E.; Novembre, J.; & Wayne, R. 2010. Genome-wide SNP and haplotype analyses reveal a rich history underlying dog domestication. Nature 464: 898-902.

Wang, G.; Zhai, W.; Yang, H.; Wang, L.; Zhong, L.; Liu, Y.; Fan, R.; Yin, T.; Zhu, C.; Poyarkov, A.; Irwin, D.; Hytönen, M.; Lohi, H.; Wu, C.; Savolainen, P.; Zhang, Y. 2016. Out of southern East Asia: the natural history of domestic dogs across the world. Cell research 26: 21-33.

Weedon, M.; Lettre, G.; Freathy, R.; Lindgren, C.; Voight, B.; Perry, J.; Elliot, K.; Hackett, R.; Guiducci, C.; Shields, B.; Zeggeni, E.; Lango, H.; Lyssenko, v.; Timpson, N.; Burtt, N.; Rayner, N.; Saxena, R.; Ardlie, K.; Tobias, J.; Ness, A.; Ring, S.; Palmer, C.; Morris, A.; Peltonen, L.; Salomaa, V.; The Diabetes Genetics Iniative, The Wellcome Trust Case Control Consortium,Smith, G.; Groop, L.; Hattersley, A.; McCarthy, M.; Hirschhorn, L.; Frayling, T. 2007. A common variant of HMGA2 is associated with adult and childhood height in the general population. Nature Genetics 39: 1245-1250.

The Westminster Kennel Club. http://www.westminsterkennelclub.org/. Referred 03.02.2016